

Ehren, Melanie; Scheerens, Jaap

Evidenzbasierte Referenzrahmen zur Schulqualität als Grundlage von Schulinspektion

Pietsch, Marcus [Hrsg.]; Scholand, Barbara [Hrsg.]; Schulte, Claudia [Hrsg.]: *Schulinspektion in Hamburg. Der erste Zyklus 2007 - 2013: Grundlagen, Befunde und Perspektiven*. Münster [u.a.] : Waxmann 2015, S. 233-272. - (HANSE - Hamburger Schriften zur Qualität im Bildungswesen; 15)



Quellenangabe/ Reference:

Ehren, Melanie; Scheerens, Jaap: Evidenzbasierte Referenzrahmen zur Schulqualität als Grundlage von Schulinspektion - In: Pietsch, Marcus [Hrsg.]; Scholand, Barbara [Hrsg.]; Schulte, Claudia [Hrsg.]: *Schulinspektion in Hamburg. Der erste Zyklus 2007 - 2013: Grundlagen, Befunde und Perspektiven*. Münster [u.a.] : Waxmann 2015, S. 233-272 - URN: urn:nbn:de:0111-pedocs-115366 - DOI: 10.25656/01:11536

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-115366>

<https://doi.org/10.25656/01:11536>

in Kooperation mit / in cooperation with:



WAXMANN
www.waxmann.com

<http://www.waxmann.com>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Marcus Pietsch, Barbara Scholand,
Klaudia Schulte (Hrsg.)

Schulinspektion in Hamburg

Der erste Zyklus 2007–2013:
Grundlagen, Befunde,
Perspektiven



HANSE
Hamburger Schriften zur Qualität
im Bildungswesen

herausgegeben von
Detlef Fickermann und Knut Schwippert

Band 15



Waxmann 2015
Münster • New York

Marcus Pietsch, Barbara Scholand,
Klaudia Schulte (Hrsg.)

Schulinspektion in Hamburg

Der erste Zyklus 2007–2013:
Grundlagen, Befunde und Perspektiven



Das gedruckte Buch finden Sie [hier](#).



Waxmann 2015
Münster • New York

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation
in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten
sind im Internet über <http://dnb.d-nb.de> abrufbar.

HANSE – Hamburger Schriften zur Qualität im Bildungswesen, Band 15

ISSN 1864-2225

ISBN 978-3-8309-3278-9

© Waxmann Verlag GmbH, 2015
Steinfurter Straße 555, 48159 Münster

www.waxmann.com
info@waxmann.com

Umschlaggestaltung: Pleßmann Design, Ascheberg
Umschlagfoto: © Robert Kneschke – Fotolia.de
Lektorat und Satz: Judith Zimmer, Hamburg
Druck: Mediaprint, Paderborn

Gedruckt auf alterungsbeständigem Papier,
säurefrei gemäß ISO 9706

Alle Rechte vorbehalten. Nachdruck, auch auszugsweise, verboten.
Kein Teil dieses Werkes darf ohne schriftliche Genehmigung des Verlages
in irgendeiner Form reproduziert oder unter Verwendung elektronischer
Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Inhalt

Vorwort.	9
---------------	---

Teil 1: Die Grundlagen der Schulinspektion

Norbert Maritzen

Schulinspektion – Aspekte der Transformation von Governance-Strukturen . . .	13
--	----

Norbert Maritzen & Herbert Altrichter

Schulinspektion in Hamburg – ein Gespräch über Startbedingungen und Entwicklungen	37
--	----

Martina Diedrich

Aufbau und Rolle der Schulinspektion Hamburg	57
--	----

Thomas Brüsemeister

Erfahrungsbereiche der Hamburger Schulinspektion – Governance-analytische Programmtheorie	77
--	----

Susanne Ulrich

Das Qualitätsleitbild der Schulinspektion Hamburg	101
---	-----

Teil 2: Die Arbeit der Schulinspektion

Marcus Pietsch, Ann-Katrin van den Ham & Olaf Köller

Wirkungen von Schulinspektion: Ein Rahmen zur theoriegeleiteten Analyse von Schulinspektionseffekten	117
---	-----

Moritz G. Sowada

Expertenurteile – Achillesferse oder Trumpf der Schulinspektion?	137
--	-----

Knut Schwippert

Daten für die Schulentwicklung – auf die Perspektive kommt es an	157
--	-----

<i>Barbara Scholand unter Mitarbeit von Stephanie Graw</i>	
Wie Schulinspektor/innen fragen und Grundschüler/innen antworten	177

<i>Peter Schulze & Christiane von Schachtmeyer</i>	
Der Blick aus der Praxis – die Schulformexpertinnen und -experten	219

Teil 3: Die Befunde der Schulinspektion

<i>Melanie Ehren & Jaap Scheerens</i>	
Evidenzbasierte Referenzrahmen zur Schulqualität als Grundlage von Schulinspektion	233

<i>Holger Gärtner</i>	
Zusammenhang von Schul- und Unterrichtsqualität und schulischen Rahmenbedingungen	273

<i>Marcus Pietsch & Timo Ehmke</i>	
Die Qualität des Unterrichts an SINUS-Grundschulen. Eine Programmevaluation mithilfe von Daten der Schulinspektion Hamburg	295

<i>Klaudia Schulte & Markus Lücken</i>	
Der Einfluss schulischer Prozesse auf die Lernentwicklung der Schülerinnen und Schüler an weiterführenden Schulen in Hamburg	317

<i>Susanne Peters</i>	
Eltern als Stakeholder von Schule. Erkenntnisse über die Sichtweise von Eltern durch die Hamburger Schulinspektion.	341

Teil 4: Die Zukunft der Schulinspektion

<i>Sabine Müller</i>	
Zukunft der externen Evaluation – Trends und Perspektiven in Deutschland. .	369

<i>Hans-Jürgen Kuhn</i>	
Rechtliche, strukturelle und politische Rahmenbedingungen für externe Evaluation in Deutschland.	385

Martina Diedrich

Der zweite Zyklus der Schulinspektion Hamburg: ein Ausblick. 419

*Martin Spiewak, Angela Hellberg, Claudia Lemke, Gerrit Petrich,
Norbert Rosenboom, Mina Zulal*

Die Perspektive der Stakeholder: Ein Interview von Martin Spiewak 437

Autorinnen und Autoren 455

Vorwort

Vor rund zehn Jahren wurden die ersten Schulinspektorate in Deutschland eingeführt. Als Teil einer veränderten Steuerungsstrategie, die die Qualitätssicherung und -entwicklung von Schulen ins Zentrum setzt, sollen sie die Qualität schulischer Prozesse evaluieren, prozessuale Mindeststandards in Schule und Unterricht sichern und darüber hinaus Schulentwicklung stimulieren.

Die Hamburger Schulinspektion ist keine Inspektion der allerersten Stunde, wurde sie doch erst im Jahr 2007 offiziell eingerichtet. Zu diesem Zeitpunkt waren in vielen anderen Bundesländern schulische Evaluatorinnen und Evaluatoren bereits zwei bis drei Jahre im Feld und hatten die unterschiedlichsten Verfahren entwickelt und verschiedene Konzepte erprobt. Von ihren vielfältigen Erfahrungen mit diesem in Deutschland dahin unbekannten Instrument konnte die Hamburger Schulinspektion profitieren.

Gleichwohl hatte man in Hamburg bereits seit der Jahrtausendwende erste Konzepte zur externen Prozessevaluation an Einzelschulen in Hamburg diskutiert und entwickelt. Den Impuls hierfür gaben insbesondere die Ergebnisse der Längsschnittstudie LAU (Aspekte der Lernausgangslage und der Lernentwicklung), die seit Mitte der 1990er Jahre regelmäßig durchgeführt wurde. Es vergingen jedoch noch mehrere Jahre, bis die erste Hamburger Schule tatsächlich durch ein Inspektionssteam evaluiert wurde. Seitdem bewegt sich die Schulinspektion im Hamburger Schulsystem stets im Spannungsfeld von Administration, Praxis und Wissenschaft und muss dabei ihren eigenen sowie den Ansprüchen der verschiedenen schulischen und außerschulischen Stakeholder gerecht werden.

In ihrem ersten Zyklus, der im Jahr 2007 begann und bis zum Jahr 2013 andauerte, inspizierte die Schulinspektion 375 Hamburger Schulen. Während dieser sechs Jahre nahmen etwa 10.000 Lehrerinnen und Lehrer, gut 36.000 Eltern, rund 92.000 Schülerinnen und Schüler sowie etwa 2.000 Vertreterinnen und Vertreter Hamburger Unternehmen an den schriftlichen Befragungen der Inspektion teil. Darüber hinaus wurden fast 30.000 Unterrichtssequenzen durch die Inspektorinnen und Inspektoren beobachtet und etwa 2.000 Interviews mit Schulleitungen, Eltern- und Schülerräten sowie Lehrerkollegien und weiteren Mitarbeiterinnen und Mitarbeitern der inspizierten Schulen, aber auch mit Vertreterinnen und Vertretern von Ausbildungsbetrieben geführt.

Die Stärke des vorliegenden Bandes ist es, so meinen wir, dass er der besonderen Stellung der Schulinspektion im Gesamtgefüge des Hamburger Bildungssystems Rechnung trägt, indem er Akteure aus den verschiedenen Bereichen und mit unterschiedlichen Perspektiven zu Wort kommen lässt. Ein Teil der Beiträge ist empirisch ausgerichtet und analysiert die Arbeit der Inspektion oder nutzt die im Rahmen der Schulinspektion Hamburg erhobenen Daten für weiterführende

Analysen zum Hamburger Schulsystem. Andere Beiträge wiederum setzen sich mit den Rahmenbedingungen, Grundlagen und Folgen von Schulinspektion auseinander oder widmen sich den Erfahrungen und Erwartungen einzelner Akteure und Akteursgruppen.

Wir bedanken uns herzlich bei allen Autorinnen und Autoren dieses Bandes für die kooperative, kollegiale und anregende Zusammenarbeit. Insbesondere haben wir uns darüber gefreut, dass die vielen Kolleginnen und Kollegen aus diversen Kontexten und Arbeitsfeldern umstandslos bereit waren, ihre Expertise in Form fundierter und kritischer Beiträge sowie Gutachten beizusteuern. Sehr glücklich sind wir vor allem auch darüber, dass sich die Vertreterinnen und Vertreter der schulischen Stakeholder aktiv in diese Publikation eingebracht haben.

In der Zusammenarbeit mit all diesen verschiedenen Personen ist uns deutlich geworden, dass es institutionen- und akteursübergreifend ein großes Interesse daran gibt, über gute Bildung in und für Hamburg ins Gespräch zu kommen und sie gemeinsam zu gestalten. Dafür bieten – nicht nur, aber auch – die Schulinspektion und ihre Ergebnisse konstruktive Gesprächsanlässe. In diesem Sinne wünschen wir Ihnen eine ertragreiche Lektüre!

Hamburg im Juni 2015

Marcus Pietsch, Barbara Scholand & Klaudia Schulte

Evidenzbasierte Referenzrahmen zur Schulqualität als Grundlage von Schulinspektion

Melanie Ehren & Jaap Scheerens¹

Zusammenfassung

Dieser Beitrag beschreibt, wie Schulinspektorate unterschiedliche Ziele (Kontrolle, Entwicklung und Verbesserung, Verbindung zwischen verschiedenen Akteuren im Schulsystem) im konzeptionellen Rahmen der Inspektionen und ihrer Indikatoren operationalisieren. Es bietet einen Überblick über die Indikatoren, die in einer Reihe von Ländern eingesetzt werden, und zeigt beispielhaft, wie sie in den konzeptionellen Rahmen von Inspektionen einfließen und die Schulen mit dem Ziel der Kontrolle, Entwicklung und Verbindung bewerten und beurteilen. Wir beschreiben die Inspektion und Bewertung der Prozesse und Ergebnisse schulischer Bildung (wozu auch gehört, Expertenmeinungen zu formulieren) und vergleichen sie mit Inspektionskonzepten, deren Hauptaugenmerk auf der Kontrolle von Input-Erfordernissen und der Einhaltung von gesetzlichen Vorschriften liegt. Außerdem erörtern wir, wie sinnvoll und angemessen unterschiedliche Konzepte im Hinblick auf aktuelle Ergebnisse der Schulwirksamkeitsforschung sind.

Schulinspektionen können unterschiedliche Ziele verfolgen, z. B. Kontrolle, Unterstützung und die Verbindung zwischen Lehrerinnen und Lehrern, Schulen und den zuständigen Stellen im Bildungssystem. *Kontrolle* knüpft an die ursprüngliche Bedeutung des Begriffs „Inspektion“ an und steht im Zentrum einer Überwachung der Einhaltung gesetzlicher Vorgaben. In vielen Ländern gilt die Kontrolle als die wesentliche Aufgabe von Schulinspektorinnen und -inspektoren. Im Mittelpunkt steht die Überprüfung der gesetzlich vorgeschriebenen Erfordernisse, Regeln und Pflichten. *Unterstützung* geht mit dem Ziel der Entwicklung einher, sie konzentriert sich auf eine Bewertung der Schulqualität, gibt Rückmeldungen an die Schule, was ihre Stärken und Schwächen betrifft, und macht Vorschläge, wie sich die Schwächen beseitigen lassen. Kontrolle und Unterstützung und regelmäßige Schulbesuche haben oftmals Auswirkungen auf die dritte Funktion der *Verbindung* von Schulinspektionen, bei der sie als Bindeglied zwischen der Spitze des Bildungssystems (wo Normen und Regeln festgelegt werden) und den Schulen selbst fungieren (wo Bildung stattfindet und konkretisiert wird). Inspektoren und Inspektorinnen neh-

1 Aus dem Englischen übersetzt von Rita Kloosterziel.

men die Rolle der Mittler ein: Sie haben einerseits die Aufgabe, die Schulen über Beschlüsse der zentralen Stellen in Kenntnis zu setzen, und müssen andererseits die Entscheidungsträger über die Realität auf Schulebene informieren.

Laut De Grauwe (2007) spiegeln diese drei Ziele unterschiedliche Systeme im Qualitätskonzept wider, das der Inspektion zugrunde liegt, und machen deutlich, worauf jeweils das größte Gewicht gelegt wird. Die Kontrolle von Schulen beinhaltet meist eine Evaluation der Input-Indikatoren und eine Überprüfung, inwieweit die Schulen gesetzliche Vorgaben einhalten. Dagegen bewerten Schulinspektorate, die darauf abzielen, eine Entwicklung von Schulen zu fördern und anzuregen, häufig die Bildungsprozesse und den Output der Schulen. Liegt der Schwerpunkt auf dem Aspekt der Verbindung, spielen Schulinspektorate vor allem die Rolle des Vermittlers, indem sie Wissen und Informationen an die entsprechenden Stakeholder innerhalb des Schulsystems weitergeben. Evaluationen mit einem bestimmten thematischen Schwerpunkt, bei denen Schulinspektorate Informationen zu spezifischen, für die Richtlinien relevanten Themen sammeln (z. B. wie neue Richtlinien umgesetzt werden oder wie das Bildungssystem in einem bestimmten Bereich abschneidet), sind Beispiele für diese Mittlerrolle eines Schulinspektorats.

1. Kontrolle von Input, Regeln und Bestimmungen

De Grauwe (2007) und Eddy Spicer et al. (2014) stellen dar, wie Inspektions- und Monitoringsysteme den Input einer Schule in den Fokus stellen können, also etwa die Anzahl der Lehrbücher pro Schüler oder Schülerin, die Ausbildung der Lehrerinnen und Lehrer, die Klassenstärke usw. Bei solchen Systemen geht es in erster Linie darum, die Einhaltung von Regeln zu kontrollieren. Sie wollen sicherstellen, dass sich Schulen an gesetzlich festgelegte Normen halten und auf der Verwaltungsebene in Übereinstimmung mit den gültigen Verordnungen handeln. Dazu gehört, dass Leitlinien und Vorschriften etwa in Bezug auf Zugangs- oder Sicherheitsbestimmungen und immer öfter auch auf das zufriedenstellende Ausfüllen der Unterlagen zur Selbsteinschätzung durch die Schulen vorliegen und angewandt werden. Als Beispiele seien das schwedische Schulinspektorat, das kontrolliert, in welchem Maße Schulen allen Schülerinnen und Schülern den gleichen Zugang zu Bildung ermöglichen, und das niederländische und das irische Schulinspektorat genannt, die überprüfen, ob Schulen ein Minimum an Unterrichtsstunden einplanen und anbieten (vgl. Ehren et al. 2013). Nach De Grauwe (2007) entspricht diese Art der Kontrolle der ältesten bürokratischen Form des Monitorings: überprüfen, ob die Regeln und Bestimmungen respektiert werden. Das klassische Inspektoratsystem in Verbindung mit unterschiedlichen Formen administrativer Selbstberichte durch Schulen (Formulare ausfüllen) ist das wesentliche Instrument, auf das sich diese Art des Monitorings stützt.

2. Evaluation und Förderung von Bildungsprozessen

Zu Bildungsprozessen zählen die Qualität des Unterrichts an einer Schule, die Interaktion zwischen Lehrerinnen und Lehrern, Schülerinnen und Schülern und Curriculum während des Unterrichts und die administrativen Organisationsprozesse innerhalb der Schule. Als Teil des konzeptionellen Rahmens von Inspektionen werden diese Prozesse immer bedeutsamer, da man übereinstimmend davon ausgeht, dass bei dem Versuch, Unterschiede in der Qualität von Schulen zu erklären, Prozessvariablen wichtiger sind als Input-Variablen. Und Informationen über die Qualität einer Schule sind notwendig, um die Qualität von Schulen insgesamt zu verbessern. Das gilt vor allem in Ländern mit hohem Einkommen, wo sich der Input der Schulen nur geringfügig voneinander unterscheidet (siehe Hanushek 1997).

Eine Übersicht von Van Bruggen (2010) zeigt z. B., dass achtzehn europäische Schulinspektorate Indikatoren und Kriterien zu „Organisation und Management in der Schule“ und „Lehren und Lernen“ in ihren konzeptionellen Rahmen aufgenommen haben, um eine nationale Perspektive auf gute Bildung zu gewährleisten und um Schulen vor dem Hintergrund eines gemeinsamen Kriterienbündels zu beurteilen, das eine solche Perspektive auf gute Bildung repräsentiert. Laut Ehren et al. (2013) orientieren sich viele dieser Konzepte an Schulwirksamkeitsforschungen. Ihre vergleichende Studie zu Inspektionskonzepten aus sechs europäischen Ländern belegt eine starke Konzentration auf Bildungsprozesse wie die Lernmöglichkeiten und Lernzeit, Leistungsorientierung, klare und gut strukturierte Unterrichtsführung, anspruchsvolle didaktische Ansätze und eine geordnete Lernumgebung. Diese Indikatoren sind in gewissem Maße Teil der Inspektionskonzepte aller sechs Länder, die Ehren et al. in ihrer Studie untersuchen (England, die Niederlande, das österreichische Bundesland Steiermark, die Tschechische Republik, Irland und Schweden).

3. Evaluation des Outputs von Schulen

Die zunehmende Verfügbarkeit und Verwendung von Daten zu Schülerleistungen haben dazu geführt, dass das Hauptaugenmerk bei Schulinspektionen immer mehr auf dem Output von Schulen liegt. Einige der Schulinspektorate (z. B. in den Niederlanden) wählen diesen Fokus auch, um die Autonomie der Schulen in Bezug auf die Gestaltung ihrer schulischen Prozesse zu respektieren. Ein verstärktes Interesse am Output von Schulen geht oft Hand in Hand mit dem allgemeinen Wunsch auf Regierungsseite, Leistungsmessungen vorzunehmen und Rankinglisten anzulegen, um auf diese Weise den Wettbewerb der Schulen untereinander zu fördern (vgl. Tolofari 2005).

Das Ziel dieser Art des Monitorings ist es, sicherzustellen, dass Schülerinnen und Schüler das lernen, was sie lernen sollen, und den Output der Schule im Zusammenhang mit Lernfortschritten der Schülerinnen und Schüler zu evaluieren. Dazu gehören z. B. die Schülerleistungen in einer Reihe von Unterrichtsfächern und die Anzahl der Schulabschlüsse bzw. Schulabbrüche. Die Evaluation von Leistungen kann auch eine Zielverschiebung verhindern, wenn sich Schulen auf die Vorlage von Vorgaben und Prozeduren konzentrieren, um die Schulinspektorate über ihre schulischen Prozesse zu unterrichten. Die Ergebnisse von Schülerleistungen in landesweit standardisierten Lernstandserhebungen und Vergleichsarbeiten werden aggregiert, um die Leistung von Schulen zu evaluieren. In manchen Ländern werden auch Rankinglisten erstellt.

In jüngster Zeit haben einige Schulinspektorate (z. B. in Norwegen, Schottland und den Niederlanden) damit begonnen, auch soziale Outcomes von Schulen zu evaluieren. Soziale Outcomes werden von Ehren und Dijkstra (2014) definiert als „der individuelle und kollektive Nutzen von Bildung bei interpersoneller Interaktion in den nicht-ökonomischen Lebensbereichen“. Auf den schulischen Bereich bezogen bestehen die sozialen Outcomes von Bildung nach Angaben der Autoren in der Kompetenz, mit anderen zusammenzuleben, und aus den sozialen Kompetenzen, die Menschen benötigen, um ihre Ziele zu erreichen und in einer Vielzahl von Situationen mit anderen zu agieren, sowohl im Beruf als auch in anderen Zusammenhängen. Zu den sozialen Outcomes von Bildung gehören auch die staatsbürgerlichen Kompetenzen, die es braucht, um einen Beitrag zu Gesellschaft, Demokratie und den gesellschaftlichen Netzwerken zu leisten, in denen Menschen leben. Soziale Outcomes werden in den konzeptionellen Rahmen von Inspektionen aufgenommen, um ein breiter gefächertes Bild vom Output einer Schule zu erlangen und zu verhindern, dass sich Schulen allein darauf beschränken, Rechnen, Lesen und Schreiben zu lehren. Man geht davon aus, dass Testergebnisse aus kognitiven Leistungstests ein unvollständiges Bild von den Kompetenzen junger Menschen zeichnen. In vielen Ländern herrscht die Ansicht vor, dass der Kanon der Kompetenzen und Fertigkeiten weiter gefasst sein sollte, wenn Schülerinnen und Schüler all das lernen sollen, was sie brauchen, um auf dem Arbeitsmarkt zu bestehen und aktive Mitglieder der Gesellschaft sein (vgl. Dijkstra et al. 2014).

Die am weitesten verbreiteten Verfahren, die bei der Evaluation des Outputs von Schulen zum Einsatz kommen, sind die regelmäßigen Messungen von Schülerleistungen in Form von standardisierten Tests und Prüfungen, in Verbindung mit der Veröffentlichung von Rankinglisten und einem systematischen (externen) Audit von Schulen. Bei den Tests geht es vor allem um kognitive Domänen wie Mathematik und Literacy, während soziale Kompetenzen häufiger mithilfe von Beobachtungen im Schulalltag gemessen werden (vgl. ebd.). Erst seit Kurzem haben einige Länder wie die Niederlande damit begonnen, landesweit standardisierte Tests zur Messung sozialer Kompetenzen einzuführen. Bisher hat nur eine

begrenzte Anzahl von Schulen diese Tests eingesetzt, sodass auch die Benchmarks und Ziele, mit denen die Schulen im Rahmen dieser Maßnahmen verglichen und evaluiert werden könnten, noch begrenzt sind.

In einigen Ländern werden neuerdings Indikatoren für den Output von Schulen verwendet, um die Inspektionsbesuche in Schulen im Rahmen von „risikobasierten Inspektionen“ gezielter zu planen. Solche risikobasierten Inspektionsmodelle fußen auf der Annahme, dass der Output einer Schule (wie er anhand von Schülerleistungen gemessen wird) ein adäquater Prädiktor für die Unterrichtsqualität dieser Schule ist und in Risikoanalysen und Früherkennung von Problemen eingesetzt werden kann, um potenziell versagende Schulen zu ermitteln. Schülerergebnisse (um die sozioökonomischen Faktoren bereinigt) in standardisierten Tests und Prüfungen werden in den Niederlanden z.B. verwendet, um Schulen in eine von drei Kategorien (grün, orange, rot) einzuordnen. Von Schulen in der grünen Kategorie nimmt man an, dass bei ihnen kein Risiko zu scheitern vorliegt. Schulen in der Kategorie Orange weisen ein potenzielles Versagensrisiko auf. Rote Schulen haben dagegen ein hohes Risiko zu versagen. Schulen in der grünen Inspektionskategorie werden lediglich einer einfachen Inspektion unterzogen, was bedeutet, dass es in dem Jahr, in dem die Kategorisierung vorgenommen wurde, keine weiteren Aktivitäten im Zusammenhang mit einer Inspektion gibt. Zusätzlich wird eine Sichtung und Analyse von Selbsteinschätzungen und anderen Dokumenten der Schule (*Desk Research*) für die Schulen im orangefarbenen Bereich angesetzt, während für Schulen in der roten Kategorie sofort Inspektionsbesuche geplant werden.

Der verstärkte Fokus auf den Output von Schulen rechtfertigt eine detailliertere und differenziertere Analyse und Verwendung von Daten zu Schülerleistungen, um die Genauigkeit der Bewertung der Inspektionsergebnisse und die Verlässlichkeit der Frühwarnanalysen zu verbessern. Aus diesem Grund haben einige Schulinspektorate unlängst damit begonnen, Messungen des *Value Added* vorzunehmen, um die Wirkung zu erfassen, den die Einzelschule und ihre Lehrerinnen und Lehrer auf die Qualität der Bildungsarbeit haben. *Value-Added*-Messungen bedienen sich mathematischer Algorithmen bei dem Versuch, den Beitrag der Schule zur Lernleistung der Schülerinnen und Schüler getrennt von allen anderen Faktoren zu betrachten, die schulische Leistungen und Fortschritte beeinflussen können, z.B. Fähigkeiten des einzelnen Schülers, Höhe des Familieneinkommens, Bildungshintergrund der Eltern oder der Einfluss von Peergroups. Das Ziel dieses *Value-Added*-Ansatzes ist es, „die Auswirkungen einzelner Lehrerinnen und Lehrer oder Schulen auf Schülerleistungen zu erfassen und dabei die Unterschiede im sozialen und familiären Hintergrund der Schülerinnen und Schüler in Betracht zu ziehen“ (ASA 2014, S. 1; Übersetzung: R.K.).

Hamilton und Koretz (2002, S. 23) unterscheiden zwischen zwei Möglichkeiten, Testergebnisse so darzustellen, dass die Leistung der Schule nachvollziehbar wird (wenn sie nach Schultyp aggregiert wird). In normorientierten Berichten wird

die Leistung einer Schule in Bezug auf ihre Stellung innerhalb einer Verteilung von Punktwerten anderer Schulen beschrieben. Solche Berichte können folgende Basis haben:

1. *Prozentrang*, der den prozentualen Anteil einer Referenzgruppe anzeigt (wobei es sich oft um die nationale Schülerpopulation in Schulen handelt), die eine niedrigere Punktzahl erzielt hat als eine bestimmte Einzelschule. Eine Schule, die im nationalen Vergleich einen durchschnittlichen Prozentrang (*National Percentile Rank* – NPR) von 75 erzielt, erreicht also eine höhere Punktzahl als 75 Prozent der Schulen, die in der nationalen Stichprobe vertreten sind.
2. *Standardwert*: Stellt die Leistung einer Schule so dar, dass man sieht, wie weit das Testergebnis vom Mittelwert/Durchschnittswert entfernt ist. Die Punktzahlen werden transformiert, sodass sie einen bestimmten Mittelwert und eine Standardabweichung (SD) aufweisen. Beispiele sind z-Werte (Mittel = 0, SD = 1), T-Werte (Mittel = 50, SD = 10) und *Normal Curve Equivalents* (NCE) (Mittel = 50, SD = 21,06). Eine Schule mit einem T-Wert von 60 liegt also um eine Standardabweichung über dem Mittel, was ungefähr einem Prozentrang von 84 entspricht.
3. *Klassenstufenäquivalente* stellen dar, wie die Leistung eines Schülers oder einer Schülergruppe im Hinblick auf die Klassenstufe einzuordnen ist, für die eine solche Leistung typisch wäre. Die Ergebnisse werden meist als Dezimalzahl dargestellt, z. B. 5,7. Dabei steht die Zahl vor dem Komma für die Klassenstufe und die Zahl hinter dem Komma für den Monat (wenn man von einem zehnmonatigen Schuljahr ausgeht, steht die Null für die Leistung der Schülerinnen und Schüler, die ganz am Anfang dieser Klassenstufe stehen). Ein Schüler oder eine Schülerin, der oder die einen Wert von 5,7 in einem Test für Viertklässler erreicht, weist dasselbe Leistungsniveau auf wie ein durchschnittlicher Schüler und oder eine Schülerin im siebten Monat der fünften Klasse, der denselben Test absolviert. Klassenstufenäquivalente stellen eine *Entwicklungsskala* dar, mit deren Hilfe Wachstum untersucht werden soll. Der mittlere Leistungszuwachs im Laufe eines Wachstumsjahres beträgt in Fach und Stufe grundsätzlich 1,0 Klassenstufenäquivalent.

Die Alternative zum normorientierten Bericht (aggregierter) Testwerte ist die kriteriums- oder lehrzielorientierte Darstellung. Bei dieser Art des Berichts geht es nicht um einen Vergleich mit anderen Schulen oder Gruppen von Schulen. Vielmehr wird die Schulleistung mit einem oder mehreren zuvor definierten Leistungsniveaus verglichen. Solche festgelegten Leistungsniveaus schließen normalerweise Zielvorstellungen im Hinblick auf minimale Testwerte und Stoffe ein, die die Schülerinnen und Schüler in spezifischen Fächern und Bereichen bewältigen können sollen.

Bei der Zielsetzung unterscheiden Hamilton und Koretz (2002) zwischen zwei breiten Ansätzen: Statusmessungen und Veränderungsmessungen. Die Statusmessung vergleicht die Leistung einer Einheit zu einem bestimmten Zeitpunkt mit einem einzigen Standard, wobei es sich um ein vom Schulinspektorat vorgegebenes Leistungskriterium wie die durchschnittliche Leistung vergleichbarer Schulen (z. B. mit ähnlichen Schülerpopulationen) oder einen historischen Durchschnitt (z. B. den Durchschnitt einer Gruppe von Schulen über einen Zeitraum von fünf Jahren hinweg) handeln kann.

Dagegen vergleichen Veränderungsmessungen die Leistung einer Einheit zu einem bestimmten Zeitpunkt mit der Messung einer früheren Leistung. Veränderung lässt sich durch einen Querschnittansatz messen, bei dem die aktuellen Viertklässler mit den Viertklässlern aus dem Jahr zuvor verglichen werden. Man kann auch einen quasi-längsschnittlichen Ansatz wählen, bei dem die aktuellen Viertklässler mit den Drittklässlern aus dem Jahr zuvor verglichen werden. Daneben gibt es noch den Längsschnittansatz, bei dem die Leistung einzelner Schülerinnen und Schüler dazu genutzt wird, Schülerinnen und Schüler über einen bestimmten Zeitraum hinweg mit sich selbst zu vergleichen. Die Ziele quantifizieren das Ausmaß der Veränderung, das man von den Schulen erwartet. Nach Hamilton und Koretz (2002) lassen sich Testwerte auf der Schul-, Klassen- und Fächerebene sowie auf der Ebene spezifischer Schülergruppen darstellen. Die Entscheidung, ob die Werte der Schulebene, der Klassen-/Fächerebene oder der Schülerebene dargestellt und ob für spezifische Gruppen disaggregiert werden soll, sollte nach Angaben dieser Autoren von der Überlegung abhängen, für welche Zwecke die Testergebnisse gedacht sind und welche Informationen die Stakeholder wünschen. Bei jedem dieser Modelle lassen sich die Testwerte an eine Reihe von Merkmalen von Schulen und Schülerinnen/Schülern anpassen, wobei die enge Beziehung zwischen Schülerleistungen und soziökonomischem Status sowie anderen Aspekten des Hintergrunds der Schülerinnen und Schüler in Betracht gezogen wird.

Beispiele für Schulinspektorate, die *Value-Added*-Messungen verwenden, lassen sich in den Niederlanden finden, wo Schülerleistungsdaten in separaten Leistungsgruppen auf der Grundlage des Grads der Benachteiligung klassifiziert werden (der sich meist am Bildungsstand der Eltern bemisst). Diese Klassifizierung wird benutzt, um den Output von Schulen zahlenmäßig zu bestimmen und einzustufen. Sie fließt auch in die Frühwarnanalysen ein, um potenziell versagende Schulen („*Failing Schools*“) für Inspektionsbesuche zu markieren. *Value-Added*-Messungen sind auch im konzeptionellen Rahmen von Schulinspektionen in England, Kanada und Irland vorgesehen. In Irland hat z. B. das *Department of Education and Skills* eine Strategie zur Verbesserung der Leistungen im Lesen, Schreiben und Rechnen entworfen. Der Entwurf empfiehlt die Verwendung eines Analysewerkzeugs für Benchmark-Daten, das „*Schools Like Ours*“ (etwa: „Schu-

len wie unsere“) genannt wird. Es soll einer Einzelschule nicht nur den Zugang zu ihren eigenen Daten ermöglichen, sondern auch zu denen vergleichbarer Schulen. In Kanada hat das *Literacy and Numeracy Secretariat* des Bildungsministeriums von Ontario ein Benchmark-Modul entwickelt, das ebenfalls mit „*Schools Like Ours*“ betitelt ist und bei dem mit beliebigen Kombinationen der verfügbaren Indikatoren gearbeitet wird, z. B. ähnliche demographische Voraussetzungen, aber bessere Leistungen.

Dass *Value-Added*-Indikatoren im konzeptionellen Rahmen von Schulinspektionen neuerdings mehr Beachtung finden, hängt mit der Notwendigkeit zusammen, die Beurteilung der Schulpraxis auf eine korrekte Verwendung der Ergebnisse der Schülerleistungen zu stützen. Scheerens, Glas und Thomas (2003) stellen z. B. fest, dass mehr Informationen über einzelne Schülerinnen und Schüler, Untergruppen von Schülerinnen/Schülern und allen Schülerinnen/Schülern einer Einzelschule sowie vergleichbare Daten für eine ganze Population (bzw. eine repräsentative Stichprobe) von Schulen eine zuverlässigere und aussagekräftigere Analyse von Schülerleistungsergebnissen ermöglicht (vgl. Scheerens et al. 2003).

Donaldson und Johnson (2010) geben jedoch an, dass noch Unsicherheit herrscht, wie groß der Beitrag tatsächlich ist, der Schulen beim Lernzuwachs ihrer Schülerinnen und Schüler zugeschrieben werden kann. Sie gehen davon aus, dass sich solche Modelle noch im Entwicklungsstadium befinden und daher fehleranfällig sind. Trotzdem sehen viele Schulinspektorate den Nutzen einer Entwicklung von *Value-Added*-Messungen, um die Zuverlässigkeit und Stichhaltigkeit ihrer Bewertungen zu verbessern, vor allem, wenn man sie mit ihren derzeit gebräuchlichen, eher groben Methoden vergleicht. Dabei werden die Schulleistungen mit dem durchschnittlichen Rohwert einer Population verglichen oder Angaben zu kostenlosem Schulmittagessen oder andere Daten zum sozioökonomischen Hintergrund der Schülerinnen und Schüler verwendet, um Schulen in separate ähnliche Leistungsgruppen zu klassifizieren und zu vergleichen.

4. Einsatztauglich: Die Modellierung von Schuleffektivität und drei Funktionen der Schulinspektionen

Viele Inspektionssysteme haben ihre Zielsetzung in den letzten Jahren dahingehend geändert, dass sie auf eine Verbesserung von Lehren und Lernen hinarbeiten. Dieses Ziel der Schulentwicklung hat in den letzten Jahren im Zuge der verstärkt verfolgten Strategie, Schulen autonomer und selbstbestimmter zu machen, an Bedeutung gewonnen. Einem hohen Maß an Schulautonomie steht in manchen Ländern die systematische Bewertung von Schulen gegenüber, um die Qualität und

Effektivität von Entscheidungen auf Schulniveau sicherzustellen. Nachlassende Schülerleistungsergebnisse, wie sie bei internationalen Vergleichen wie PISA und TIMSS zutage getreten sind, haben auch oft zu einem Anstieg an Beurteilungen und Kontrollen von Schulen geführt, selbst in vermeintlich dezentralisierten Bildungssystemen.

Wenn Schulinspektionen ihren Zweck erfüllen sollen, müssen die Art der Inspektion und vor allem die Standards im konzeptionellen Rahmen der Inspektion an die beabsichtigten Ziele des verbesserten Lehrens und Lernens und schlussendlich der verbesserten Schülerleistung anpasst sein. Die Fachliteratur zu Bildung und Schuleffektivität bietet wichtige Hilfestellungen bei der Definition einer „guten Schule“ und bei kritischen Überlegungen, inwieweit Inspektionsstandards in unterschiedlichen Ländern durch Forschungsergebnisse belegt sind. Ganz allgemein formuliert kann man „Bildungseffektivität“ als den Grad der Zielerreichung eines Bildungssystems verstehen. Ein Bildungssystem mag ein landesweites Bildungssystem sein, eine Einzelschule, eine Gruppe von Schülerinnen und Schülern oder sogar ein einzelner Schüler oder eine einzelne Schülerin. Im Zusammenhang mit Schulinspektionen konzentrieren wir uns hier auf Schulen und Schuleffektivitätsforschung als Kernthema.

Schuleffektivitätsforschung versucht, sich mithilfe wissenschaftlicher Methoden mit den kausalen Aspekten zu beschäftigen, die dem Konzept der Effektivität innewohnen. Anhand des Outcomes wird eine Einschätzung der Wirksamkeit von Schule vorgenommen, vor allem aber wird die Tatsache beleuchtet, dass Unterschiede in der Wirksamkeit von Schule beeinflussbaren Voraussetzungen zugeschrieben werden, zu denen sowohl Input als auch Prozesse gehören. Normalerweise ist die Einschätzung von Schulwirkung das Ergebnis von Vergleichen, wenn z. B. die durchschnittlichen Leistungswerte der Schulen miteinander verglichen werden. Leistungswerte in Kernfächern, die am Ende eines festgelegten Programms erfasst werden, stellen die wahrscheinlichsten „Schulwirkungen“ dar, obwohl man auch alternative Kriterien heranziehen kann, etwa die Zufriedenheit der Lehrerinnen und Lehrer oder die Art, wie die Schule mit ihrer Umgebung interagiert. Um die „Nettowirkung“ beeinflussbarer Bedingungen wie den Einsatz verschiedener Lehrmethoden oder eine bestimmte Art des Schulmanagements zu erfassen, bedarf es einer Anpassung der Leistungsmessungen, sodass sie die unterschiedlichen Aufnahmepraktiken von Schulen berücksichtigen. Zu diesem Zweck werden Merkmale des Hintergrunds der Schülerinnen und Schüler wie der sozioökonomische Status, die allgemeine Lernfähigkeit oder die Ausgangsleistungen in einem Fach als Kontrollvariablen eingesetzt. Diese Art der statistischen Anpassung, wie sie in Forschungsarbeiten vorgenommen wird, findet eine praktische Entsprechung in dem Bemühen, „faire Vergleiche“ zwischen Schulen anzustellen und den *Value Added* zu bemessen (siehe oben, vgl. Scheerens 2013, S. 4).

Was Schulwirksamkeitsforschung und die Wissensgrundlage, die sich aus dieser Forschungsrichtung für die Schulinspektion an sich ergeben hat, miteinander verbindet, ist in erster Linie ein Interesse am Outcome von Schule als dem ultimativen Qualitätsstandard. Für den Bereich der Schulinspektion ist die Schulwirksamkeitsforschung jedoch vor allem aus einem Grund besonders wichtig: Sie liefert empirisch gestützte Informationen zu der Frage, welche beeinflussbaren schulischen Bedingungen den größten Einfluss auf den Outcome von Schule haben, und bietet so eine wissenschaftlich fundierte Begründung für die Wahl der Input- und Prozessindikatoren.

Wenn wir von den drei grundlegenden Funktionen von Schulinspektionen ausgehen, die wir zu Beginn dieses Kapitels angeführt haben, so ist es vor allem der Bereich der Entwicklung und Verbesserung, der von einer Verknüpfung von Schulwirksamkeit und Schulinspektionen profitiert. Wenn Schulinspektionen nicht nur über die Wirkungen von Schule berichten, sondern auch über die Input- und Prozessindikatoren, könnte man erwarten, dass ein Feedback zu diesen Indikatoren unmittelbare Ansatzpunkte für Handlungen bietet, die zu einer Schulentwicklung führen. Wenn z. B. ein Schulprozessindikator wie die Verbindung zwischen Schulcurriculum und Aufgaben oder Prüfungen (oft als „Lerngelegenheiten“ angegeben) einen niedrigen Wert aufweist, wäre eine Optimierung der Übereinstimmung von Unterrichtsinhalt und Aufgabeninhalt eine plausible Vorgehensweise, um die Schulleistung zu steigern.

Beim Monitoring als einer der Funktionen von Schulinspektionen ist die Verbindung zur Schulwirksamkeit komplexer. Erstens gibt es keine Verbindung zwischen der Überwachung der Einhaltung von Regeln und der Wirksamkeit von Schule, soweit es die Erfüllung von standardmäßigen Vorschriften und Bestimmungen betrifft. Einhaltungskontrollen passen jedoch zum Verständnis von Wirksamkeit, wenn grundlegende Inputs in Schulen evaluiert werden, die sich unmittelbar auf den Outcome von Bildung auswirken. Solche Inputs sind leicht zu benennen, z. B. Lehrerqualifikationen, zahlenmäßiges Schüler-Lehrer-Verhältnis und festgeschriebene Unterrichtszeiten. Inwieweit solche Inputmessungen einen Unterschied ergeben, hängt unter anderem vom Kontext ab. Meist zeigen solche Inputs in Entwicklungsländern deutlichere Auswirkungen als in industrialisierten Ländern, da in industrialisierten Ländern normalerweise alle Schulen über diese grundlegenden Inputs verfügen und die Schulen sich in Bezug auf Lehrerqualifikationen oder reguläre Unterrichtszeiten kaum voneinander unterscheiden (vgl. Hanushek 1997). So stellt Hanushek (1986, S. 1161) fest, dass die Variable „Berufserfahrung der Lehrerinnen und Lehrer“ als einzige unabhängig vom Kontext einen zuverlässigen Einfluss auf den Outcome von Schule hat: 30 Prozent der geschätzten Koeffizienten sind statistisch signifikant. Hanushek kommt zu dem Schluss, dass Ausgaben im Bildungsbereich anscheinend nicht durchgehend im Zusammenhang mit Leistung stehen

und dass wichtige Effekte erst bei größeren Unterschieden in den Inputs zu erwarten seien.

Was die Verbindungsfunktion von Schulinspektionen angeht, so besteht lediglich ein eher theoretischer Zusammenhang zu Schulwirksamkeit und Schulinspektionen, und zwar insofern, als sich Mehrebenen-Modelle von Schulwirksamkeit vorrangig mit einer Angleichung von Facetten und Elementen befassen, die auf verschiedenen Ebenen operieren (siehe z.B. Scheerens 2007; Creemers und Kyriakides 2008). Wenn man Bildungssysteme als hierarchisch geordnete Systeme betrachtet, kann man Schulwirksamkeit von Unterrichtswirksamkeit und von Systemeffektivität unterscheiden. Dieser letztgenannte Begriff ist nicht besonders weit verbreitet, er gehört zu einer neueren Forschungsrichtung, die ihre Impulse vor allem aus dem starken Anstieg internationaler Schulleistungsuntersuchungen bezieht. Bei solchen Untersuchungen können von politischen Entscheidungen beeinflussbare Bedingungen im Rahmen des nationalen Systems mit Schülerleistungen in Verbindung gebracht werden. Beispiele für solche Entscheidungen sind Regelungen zur Förderung der Autonomie, *Accountability* und Wahlmöglichkeiten. Unterrichtswirksamkeit konzentriert sich auf effektives Lehrerhandeln im Rahmen des Unterrichts. Diese Strömung der Bildungsforschung zeichnet sich dadurch aus, dass Prozessmerkmale von Bildung auf Lehrer- oder Unterrichtsebene untersucht werden. Wenn wir uns also mit Variablen dieser Ebene beschäftigen, bei denen sich eine Verbindung zu Leistung gezeigt hat, tauchen wir in die primären Prozesse schulischer Ausbildung ein. Schulwirksamkeit richtet ihr Augenmerk dagegen auf die Voraussetzungen, die eine effektive Schule mitbringt, und befasst sich mit der Frage, welche Bedingungen für einen Zuwachs bei Schülerleistungen (*Value Added*) sorgen. Das Ziel besteht im Allgemeinen darin, Schulmerkmale aufzuspüren, die sich ganz sicher mit dem Output einer Schule in Verbindung bringen lassen. Dieser Output wird meist anhand von Schülerleistungen gemessen.

Schulwirksamkeit meint die Verschmelzung der Wirksamkeitsforschungen auf diesen drei Ebenen. Konzeptionelle Beiträge zu diesem Forschungszweig sprechen von Schulen als Ansammlungen von „*Nested Layers*“ (Purkey und Smith 1983), ineinander geschachtelten Schichten. Ihnen liegt die zentrale Annahme zugrunde, dass höhere organisatorische Levels effektivitätsfördernde Bedingungen auf niedrigeren Levels begünstigen (vgl. Scheerens und Creemers 1989). Die Mehrebenenanalyse hat einen wesentlichen Beitrag zur Entwicklung solcher integrierten Schulwirksamkeitsmodelle geleistet (vgl. Scheerens 2013).

Diese Vorstellung von Schulwirksamkeit und ineinander geschachtelten Klasebenen innerhalb der Schulebene, die wiederum Teil eines nationalen Bildungssystems ist, knüpft an die Verbindungsfunktion von Schulinspektionen an. Sie greift die besonderen Verbindungen zwischen den verschiedenen Levels auf, die Inspektionen bei ihrer Bewertung von Unterricht, Schulen und Bildungssystem

und mit ihren Informationen für die Akteure auf diesen Ebenen schaffen können. Schulinspektorate könnten die Abstimmung und Verknüpfung der ineinander geschachtelten Ebenen fördern, indem sie z. B. eine Kohärenz nationaler curricularer Konzepte mit Evaluation, Inspektion und Beurteilungskonzepten anstoßen. Theoretische Modelle zu Bildungseffektivität erkennen jedoch auch an, dass es bei vielen Bildungssystemen lediglich eine lose Kopplung zwischen den Schichten des Bildungssystems gibt. Außerdem herrschen in den meisten Ländern verschiedene Erwartungen vor, wie weit nationale Inspektorate als Instrumente der Zentralverwaltung fungieren und eine enge Kopplung fördern sollen. In manchen Fällen geht man davon aus, dass Inspektorate unabhängig oder nahezu unabhängig funktionieren und bei der Evaluation von Schulen ihrer eigenen Agenda folgen. Soweit diese Art der systemischen Angleichung bis jetzt erforscht wurde (z. B. Mourshed et al. 2010; Scheerens et al. 2015), lässt sich feststellen, dass die Funktionsweise von Inspektoraten bisher nicht thematisiert wurde.

5. Identifizierung von effektivitätsfördernden Schulvoraussetzungen; Übereinstimmung in Reviews

Der wesentliche Teil der Schuleffektivitätsforschung besteht in der Identifizierung von effektivitäts- und verbesserungsorientierten Bedingungen. Im folgenden Abschnitt werden neuere und ältere Forschungsreviews genannt und auf Übereinstimmungen in Bezug auf die wichtigsten effektivitätsfördernden Voraussetzungen untersucht. Eine solche Zusammenfassung ermöglicht den Vergleich der grundlegenden Forschungen zur Schuleffektivität mit Inspektionskonzepten in einem späteren Abschnitt.

Scheerens (2014) fasst die Ergebnisse von Review-Studien aus den 1990er Jahren und Review-Studien jüngerer Datums zusammen (z. B. Reynolds et al. 2014; Muijs et al. 2014 und Hopkins et al. 2014). Bei den älteren Reviews handelt es sich um die von Purkey und Smith (1983), Scheerens (1992), Levine und Lezotte (1990), Sammons et al. (1995) und Cotton (1995). Diese früheren Review-Studien nennen die folgenden Bedingungen als zuträglich für hohe Schülerleistungen:

- Leistungsorientierung und hohe Erwartungen: ein produktives Schulklima, eine Schulphilosophie, die auf Leistung abzielt, gemeinsame Vision und Ziele, hohe Erwartungen in die Leistungsfähigkeit aller Schülerinnen und Schüler
- Kooperative Atmosphäre und ein geordnetes Klima: kooperative Planung, Übereinstimmung in Bezug auf lernorientierte Atmosphäre, geordnetes Klima
- Klare Zielsetzung in Bezug auf Grundfertigkeiten: Fokus auf Schülerlernen, Konzentration auf Unterricht

- Häufige Evaluation: angemessenes Monitoring, evaluatives Potenzial der Schule, Beurteilung
- Berufliche Weiterentwicklung: Personalentwicklung, berufliche Fortbildungen, eine lernende Organisation
- Einbeziehung der Eltern: Unterstützung der Eltern, Zusammenarbeit von Schule und Familie
- Schulleitung: Schulleitungshandeln, Schulmanagement und -organisation, entwicklungsorientierte Schulleitung
- Effektive Unterrichtsplanung: Klassenführung, *Time-on-Task*, strukturierter Unterricht, Lerngelegenheiten, Koordination von Lehrplan und Unterricht

Die folgenden Faktoren werden von den Autoren der früheren Review-Studien übereinstimmend als die wesentlichen Unterrichtsbedingungen eingestuft: Leistungsorientierung (die eng mit „hohen Erwartungen“ verbunden ist), Kooperation, Schulleitungshandeln, häufige Evaluation, Zeit, Lerngelegenheit und „Struktur“.

Neuere Reviews von Reynolds et al. (2014), Muijs et al. (2014) und Hopkins et al. (2014) bieten einen Überblick über die wichtigsten Voraussetzungen in der Schulwirksamkeitsforschung (*Educational Effectiveness Research – EER*) und der Unterrichtswirksamkeitsforschung (*Teaching Effectiveness Research – TE*). Das Review von Hopkins et al. (2014) fügt dieser Forschung eine Komponente der Verbesserung und Entwicklung hinzu (*School System Improvement – SSI*) und zielt darauf ab, unser Verständnis effektiver Interventions- oder Entwicklungsprogramme und der Voraussetzungen in Schulen zu fördern, die zu einer wirksamen Schulentwicklung beitragen. Die Zusammenfassung dieser Studien in Tabelle 1 (zitiert nach Scheerens 2014) verdeutlicht, dass es über die Jahre einen Konsens hinsichtlich der wesentlichen Bedingungen von Schule und Unterricht gibt. Die fünf Faktoren, zu denen in den früheren Reviews die größte Übereinstimmung auszumachen war, tauchen auch in diesen neueren Arbeiten auf. Die wichtigste Entwicklung zeigt sich im Review zu Unterrichtswirksamkeit in der Form von neu hinzugekommenen, vom Konstruktivismus inspirierten Unterrichtsstrategien. Sie sind in Tabelle 1 in der zweiten Spalte kursiv gesetzt.

Tabelle 1: Effektivitätsfördernde Bedingungen, Zusammenstellung aus den Review-Studien von Reynolds et al. (2014); Muijs et al. (2014) und Hopkins et al. (2014)

EER	TE	SSI
<ul style="list-style-type: none">• Effektive Führung• Akademischer Fokus• Positives, geordnetes Klima• Hohe Erwartungen• Beobachtung der Lernfortschritte• Einbeziehung der Eltern• Effektiver Unterricht (Zeit)• Mitarbeiterfortbildung• Schülerbeteiligung	<ul style="list-style-type: none">• Innerschulische Lerngelegenheiten (<i>Opportunity to learn</i>)• Zeit• Klassenmanagement• Struktur und unterstützende Lerngerüste, u. a. Feedback• Produktives Klassenklima• Klarheit der Präsentation• <i>Selbstgesteuertes Lernen fördern</i>• <i>Vermittlung meta-kognitiver Techniken</i>• <i>Vermittlung von Modellbildung</i>• <i>Differenziertere Diagnose</i>• <i>Bedeutung von Vorwissen</i>	<ul style="list-style-type: none">• Aspekte organisatorischer Gesundheit• Schulinterne Reflexion und Verbesserungsstrategien• Schulentwicklungsplan• Umfassende Schulreform• Facetten des Schulleitungshandelns (transformativ, pädagogisch, dezentralisiert)• Effektive systemische Reform, u. a. Schülerleistungen und Betonung auf Qualität des Unterrichts

EER – Schulwirksamkeitsforschung (*Educational Effectiveness Research*); TE – Unterrichtswirksamkeitsforschung (*Teaching Effectiveness Research*); SSI – Verbesserung und Entwicklung von Schule (*School System Improvement*). Quelle: Scheerens 2014

6. Weniger Konstanz der Effektgrößen in Synthesen quantitativer Forschung

Der vorangegangene Abschnitt zeigte eine deutliche Überschneidung bei wirksamkeitsfördernden Bedingungen sowohl in qualitativen als auch in quantitativen Reviews. Die quantitativen Reviews und Metaanalysen deuten jedoch stets auf bedeutsame Unterschiede in den spezifischen Effektgrößen aller Schlüsselvariablen hin. Diese Unterschiede sind in Tabelle 2 dargestellt (zitiert nach Scheerens 2013). Die Unterschiede zwischen den von Hattie (2009) berichteten Effektgrößen einerseits und den anderen, stärker auf Europa bezogenen Metaanalysen andererseits sind recht beeindruckend. Es sollte beachtet werden, dass Hattie Effektgrößen mithilfe des d-Koeffizienten ausdrückt und die anderen Autoren Korrelationen

darstellen. Da die Korrelationskoeffizienten halb so groß sind wie d-Koeffizienten, lassen sich die Werte nährungsweise ineinander überführen.

Tabelle 2: Ergebnisse aktueller Meta-Analysen

Schulbezogene Prozessmerkmale			
	Scheerens et al. 2007	Hattie 2009	Creemers & Kyriakides 2008
Konsens/Kohärenz Kollegium	.02	–	.16
Geordnetes Klima	.13	.34	.12
Monitoring und Evaluation	.06	.64	.18
Curriculum/Lerngelegenheiten	.15	–	.15
Hausaufgaben	.07	.30	–
Effektive Lernzeit	.15	.34	–
Einbeziehung der Eltern	.09	.50	–
Leistungsorientierung	.14	–	–
Schulleitungshandeln	.05	.36	.07
Differenzierung	.02	.18	–
Unterrichtsbezogene Prozessmerkmale			
	Scheerens et al. 2007	Hattie 2009	Seidel & Shavelson 2007
Zeit und Lerngelegenheiten	.08	.34	.03
Klassenführung	.10	.52	.00
Strukturierter Unterricht	.09	.60	.02
Vermittlung von Lernstrategien	.22	.70	.22
Feedback & Monitoring	.07	.66	.01

Koeffizienten basieren auf der Fisher-z-Transformation von Korrelationen; da Hattie Effektgröße d verwendet, erscheinen diese Angaben in Fettschrift. Quelle: Scheerens 2013, S.14

Da wir auf einen Vergleich zwischen Inspektionskonzepten und Schulwirksamkeitsforschung abzielen, fügen wir die nachfolgende Tabelle mit einer durchschnittlichen Effektgröße von wesentlichen Wirksamkeitsbedingungen an, wie sie in Metaanalysen von Marzano (2003), Scheerens et al. (2007) und Hattie (2009) beschrieben werden. Eine Mittlung der Effektgrößen stellt zwar eine etwas grobe Methode dar, vermittelt aber einen Eindruck von der relativen Bedeutung dieser entschiedenen wirksamkeitsfördernden Bedingungen. Die durchschnittlichen Effektgrößen zeigen an, dass die wichtigste Bedingung des schulischen Beitrags zu guten Schülerleistungen darin besteht, die Schülerinnen und Schüler bildungsbezogenen Inhalten (Lerngelegenheiten und Unterrichtszeit) „auszusetzen“. Organisationale

Faktoren wie Schulleitungshandeln und Kooperation weisen dagegen relativ kleine Effektgrößen auf.

Tabelle 3: Rangfolge von Schulwirksamkeitsvariablen nach durchschnittlicher Effektgröße (d-Koeffizient), Zusammenstellung aus drei Meta-Analysen

	Marzano 2003	Scheerens et al. 2007	Hattie 2009	Durch- schnittliche Effektgröße
Lerngelegenheiten	.88	.30	.39*	.523
Unterrichtszeit	.39	.30	.38	.357
Monitoring	.30	.12	.64	.353
Leistungserwartungen	.27	.28	.43**	.327
Einbeziehung der Eltern	.26	.18	.50	.313
Schulklima	.22	.26	.34	.273
Schulleitungshandeln	.10	.10	.36	.187
Kooperation	.06	.04	.18***	.093

*) operationaisiert als „Zusatzprogramm für begabte Kinder“; **) operationalisiert als „Lehrererwartungen“; ***) operationalisiert als „Teamteaching“

Quelle: Scheerens 2013, S. 24

Ein weiteres wichtiges Thema, das es zu berücksichtigen gilt, ist die Konstanz dieser Effekte durch alle Einzelschulen. Konstanz in der Schätzung von Schuleffekten durch Klassenstufen, Lehrerinnen und Lehrer und Unterrichtsfächer sowie Stabilität von Schuleffekten durch Jahre stellen bei Schulinspektionen eine wichtige Grundannahme dar. Die Beurteilung der Schulqualität durch Schulinspektionen findet normalerweise einmal in drei bis fünf Jahren statt und man geht davon aus, dass die Schul- und Unterrichtsqualität bis zum nächsten Inspektionsbesuch relativ unverändert bleibt.

Mehrere Studien zur Schulwirksamkeit stellen die potenzielle Beständigkeit der Schulwirksamkeit (oder den Mangel derselben) mithilfe einer Analyse einer Korrelationsmatrix zu fächer- und kohortenbezogenen (oder klassenbezogenen) Effekten dar und berechnen die Stärke eines allgemeinen Schulfaktors (vgl. Scheerens 2013). Dazu seien Vermeer und Van der Steeg (2011) zitiert:

„Typischerweise ist die Rangordnung der durchschnittlichen Leistung (*Value Added*) von Schulen durch die Jahre korreliert. In einer Studie zu weiterführenden Schulen in den Niederlanden fanden Bosker et al. (1989) Korrelationen, die abhängig von der Zeitspanne von ein bis vier Jahren abnahmen: .74 (ein Jahr), .62 (zwei Jahre), .49 (drei Jahre) und .49 (vier Jahre). Gray et al. (1995) betrachteten bei weiterführenden Schulen in England Zeitabstände von einem, zwei und drei Jahren und

fanden Korrelationen von .94, .96 und .81. Thomas et al. (2010) analysierten in der Grafschaft Lancashire Schuldaten über einen Zeitraum von elf Jahren hinweg. Sie kamen zu dem Schluss, dass die Schuleffekte vergleichsweise stabil waren. Wenn Schulen jedoch als durchschnittlich, über- oder unterdurchschnittlich kategorisiert wurden, ließen sich in den einzelnen Kategorien viele Veränderungen feststellen. Im Verlauf von elf Jahren waren 50 Prozent der Schulen in eine andere Kategorie gewechselt. Außerdem waren kaum beständige Fortschritte zu beobachten:

Bei den meisten Schulen scheint der Aufwärtstrend typischerweise nicht länger als drei Jahre anzudauern. Kurz gesagt erbringt die nichtlineare Modellierung den Nachweis, dass es zwar zweifelsohne Veränderungen gab, dass diese jedoch nicht ‚fortdauernd‘ waren und in vielen Fällen auch zufällig aufgetreten sein könnten. Diese Ergebnisse stehen in krassem Widerspruch zu Wunschkonzeptionen der Regierung von einer stetig fortschreitenden Schulentwicklung (Thomas et al. 2010, S. 280). Auch eine neuere niederländische Studie weist eine geringere Stabilität nach: Von den weiterführenden Schulen mit den höchsten Punktwerten waren nach drei Jahren nur noch 15 Prozent in der obersten Kategorie.“ (Vermeer & Van der Steeg 2011, S. 9 f.; Übersetzung: R. K.)

Um der mangelnden Stabilität Rechnung zu tragen, wäre es sinnvoll, die Position von Schulen über eine bestimmte Zeitspanne (z. B. drei Jahre lang) im Rahmen von Rechenschafts- und Belohnungsprogrammen zu begutachten und Schulen anhand ihres Leistungsdurchschnitts aus einer Reihe von Jahren zu vergleichen. Die Forschungsergebnisse aus verschiedenen Ländern zeigen diesbezüglich wiederum, dass die Anzahl der Jahre der Durchschnittsberechnung in jedem Land unterschiedlich sein kann, da die Stabilität von Schuleffekten von Land zu Land verschieden zu sein scheint und eventuell die Homogenität von Bildungssystemen widerspiegelt.

7. Was mit den Schlüsselfaktoren schulischer Wirksamkeit gemeint ist

Um ein wenig genauer hinter die „Etiketten“ der Faktoren zu sehen, die in den Tabellen der vorangegangenen Abschnitten genannt wurden, folgt nun eine Beschreibung der wesentlichen Merkmale und der Subkomponenten dieser Faktoren.

Leistungsorientierung

Dieser Faktor bezeichnet ergebnisorientiertes Streben und eine positive, optimistische Einstellung zur Leistungsfähigkeit aller Schülerinnen und Schüler. Zu den Datenquellen gehören Planungsunterlagen wie Schulentwicklungspläne oder Leit-

bilder, Antworten von Schulleitung und Lehrerinnen und Lehrern zu Fragebögen und Nachweise auf Verwaltungsebene, z. B. Aufzeichnungen zu Schülerleistungen. Leistungsorientierung wird oft in Subkomponenten unterteilt, z. B. „klarer Fokus auf Beherrschung der Kernfächer“, „hohe Erwartungen auf Schul- und Lehrerebene“ und „Aufzeichnungen zu Schülerleistungen“.

Schulleitungshandeln

Bei vielen operationalen Definitionen und Instrumenten im Zusammenhang mit Schulleitungshandeln stellt man einen deutlichen Fokus auf Führungsrollen im Bereich des Hauptprozesses des Lernens und Unterrichtens und im Bereich organisatorischer Voraussetzungen, die als Unterstützung für diesen Hauptprozesses angesehen werden, fest. Dazu gehören das Coaching von Lehrerinnen und Lehrern und Hilfestellung bei curricularen Entscheidungen. Oft wird auch ein Zusammenhang zum Monitoring von Einschätzungen und Fortschritten von Schülerinnen und Schülern hergestellt. Neben dem Schulleitungshandeln, das sich auf den primären Prozess des Unterrichtens und Lernens konzentriert, ist „transformationales Schulleitungshandeln“ eher auf Aktivitäten gerichtet, die die Schule auf organisatorischer Ebene verbessern. „Dezentralisiertes Schulleitungshandeln“ oder gar „*Teacher Leadership*“ bedeutet, dass Teile der Schulleitung an Lehrerinnen und Lehrer abgegeben werden. Relevante Subkomponenten von Schulleitungshandeln sind: allgemeiner Führungsstil, Führungsrollen im Zusammenhang mit Koordination (partizipative Entscheidungsprozesse und Informationsvermittlung steuern), Meta-Kontrolle von Unterrichtsprozessen, Förderung von Personalfortbildung.

Kooperation, Konsens und Kohärenz im Kollegium

Früher gab es eine leichte Betonung auf messbaren Facetten von Kooperation (wie Häufigkeit von Konferenzen) und persönlicher Zufriedenheit. Neuerdings setzt sich eher der Gedanke von Schulen als professionellen Lerngemeinschaften und *Peer Learning* durch, sodass Kooperation stärker im Zusammenhang mit schulbezogenen Entwicklungsinitiativen einerseits und Diskussionen zu Unterricht und Lernen andererseits gesehen wird. „*Teamteaching*“ gilt auch oft als Indikator für Kooperation unter Kollegen.

Kooperation, Konsens und Kohärenz im Kollegium wird normalerweise durch die Art und Häufigkeit von Zusammenkünften und Besprechungen, Zufriedenheit mit Kooperation, aufgabenbezogene Facetten von Kooperation, Konstanz in Bezug auf Unterrichtsziele und -methoden gemessen.

Curriculare Qualität und Lerngelegenheiten

Curriculare Qualität wird meist im Sinne systematischer Prozessplanung und erlebter Zufriedenheit mit dem Curriculum gemessen. Das Konzept der Lerngelegenheiten steht im Zusammenhang mit der Anpassung von schulischen Zielen, Unterricht und Schülerbeurteilung. Grundsätzlich geht es um die Übereinstimmung der unterrichteten Inhalte mit den getesteten Inhalten. In jüngeren Studien wird „Testvorbereitung“ als neue Sichtweise auf Lerngelegenheiten betrachtet. Bezogen auf den Unterricht ist „unterrichtliche Angleichung“ eine weitere neue Interpretation von Lerngelegenheiten. Im Laufe der Zeit wurden curriculare Qualität und Lerngelegenheiten in folgende Subkomponenten operationalisiert: systematisches Formulieren und Überwachen curricularer Prioritäten, Auswahl und Anwendung von Methoden und Lehrbüchern und Lerngelegenheiten im Sinne von behandelten Inhalten.

Schulklima

Ein geordnetes und sicheres Schulklima ist eine der organisatorischen Voraussetzungen von Schulen, die in internationalen Schulleistungsuntersuchungen wie PISA relativ positive Unterstützung erfahren haben. Außerdem hat die Schulwirksamkeitsforschung diesen Punkt von Anfang an im Blick. Offensichtlich ist die leistungsbezogene Facette des Schulklimas eng verbunden mit „Leistungsorientierung“ und „leistungsorientierter Schulphilosophie“, wie sie oben beschrieben wurden. Interne Beziehungen als Teil des „beziehungsorientierten Schulklimas“ betreffen die Beziehungen zwischen Lehrerinnen/Lehrern und Schülerinnen/Schülern, Lehrerinnen/Lehrern und Schulleitung sowie Lehrerinnen/Lehrern und Schülerinnen/Schülern jeweils untereinander. Relevante Subkomponenten des Schulklimas sind Disziplin, Leistungsorientierung und gute interne Beziehungen der Mitarbeiter einer Schule untereinander sowie zwischen Mitarbeitern und Schülerinnen und Schülern.

Evaluation und Monitoring

Evaluation als eine wirksamkeitsfördernde Bedingung hat mit der Evaluationsorientierung (oder dem Mangel einer solchen Orientierung) an einer Schule zu tun, sowohl auf den Unterricht als auch auf die Schülerinnen und Schüler bezogen. Die Häufigkeit der Anwendung wird ebenso gemessen wie die Zufriedenheit der Mitarbeiter mit Evaluationen und die Art und Weise, wie Evaluationsergebnisse im Sinne einer Schulentwicklung eingesetzt werden. Zur Messung von Eva-

uation und Monitoring wird zunächst kontrolliert, ob eine Schule regelmäßig systematische Selbstevaluationen und Schülerbewertungen vornimmt und/oder andere Arten von Tests und regelmäßige Überprüfungen des Leistungsstandes der Schülerinnen und Schüler einsetzt. Intensivere Messungen von Evaluation und Monitoring beinhalten eine Analyse einer aufgabenbezogenen Zusammenarbeit der Lehrerinnen und Lehrer und gehen der Frage nach, ob (formell standardisierte und informelle Formen von) Lehrer- und Unterrichtsevaluationen einen festen Platz im *Peer Learning* haben. Subkomponenten von Evaluation und Monitoring sind Schulevaluation, Klassenevaluation und Überprüfung von Schülerleistungen.

Einbeziehung der Eltern

Wesentliche Bestandteile der Einbeziehung der Eltern sind das Mitspracherecht bei der Formulierung von Schulprinzipien, die aktive Beteiligung der Eltern an schulischen Angelegenheiten, die Unterstützung durch die Eltern entweder in Form praktischer Hilfestellungen oder im Zusammenhang mit Unterricht und Lernen. Die ambitionierteste Form der Zusammenarbeit mit den Eltern ist sicherlich der Versuch der Schule, das pädagogische Klima in der Familie zu beeinflussen. Die Einbeziehung der Eltern wird oft dadurch gemessen, dass Schulen befragt werden, welchen Stellenwert sie der Beteiligung von Eltern in ihrer Schulphilosophie einräumen, wie intensiv der Kontakt zu Eltern ist und wie zufrieden die entsprechenden Akteure (Lehrerinnen und Lehrer, Eltern und Schulleitung) mit der Einbeziehung der Eltern sind. Das Konzept der Einbeziehung der Eltern lässt sich zu einer „Einbeziehung des schulischen Umfelds“ erweitern.

Klassenklima

Das Klassenklima deckt sich in weiten Teilen mit der Beschreibung des Schulklimas, vor allem in Bezug auf die Subkomponenten. Beide Definitionen beinhalten einen disziplinarischen Teil und heben gute Beziehungen hervor, das Klassenklima konzentriert sich jedoch auf die Beziehung zwischen Lehrerinnen und Lehrern und Schülerinnen und Schülern, während das Schulklima auch die Beziehungen des Schulpersonals untereinander einbezieht. Zum Klassenklima gehört darüber hinaus die Idee von kognitiver und emotionaler Unterstützung von Schülerinnen und Schülern, in der Form von klaren und verständlichen Erklärungen, Hilfe bei Aufgaben sowie einer anregenden Einbeziehung und einem Gefühl der Selbstwirksamkeit der Schülerinnen und Schüler. Bisweilen wird beim Klassenklima auch ein „Spaßfaktor“ berücksichtigt: Die Schülerinnen und Schüler werden gefragt,

ob sie ihren Lehrer oder ihre Lehrerin mögen, ob er oder sie mit ihnen auch über außerschulische Aktivitäten spricht und ob sie auch Witze machen und lachen.

Effektive Lernzeit

Lernzeit kann ganzheitlich oder detaillierter gemessen werden. Eine differenzierte Messung unterscheidet zwischen zugewiesener Lernzeit (offiziell anberaumte Unterrichtsstunden), reiner Unterrichtszeit (der Teil der Stunde, den der Lehrer tatsächlich auf das Unterrichten verwendet, wobei die Zeit für Organisatorisches und für Ablenkungen abgezogen wird) und *Time-on-Task* (die Zeit, in der ein Schüler oder eine Schülerin aktiv beschäftigt ist). Klassenführung wird oft im Sinne einer Maximierung der reinen Unterrichtszeit und der *Time-on-Task* definiert. Eine weitere wichtige Unterscheidung gibt es zwischen der Unterrichtszeit in der Schule und der Zeit, die Schülerinnen und Schüler mit Hausaufgaben zubringen. Studien zur Unterrichtszeit beinhalten oft Informationen zu Abwesenheitszeiten von Schülern und Unterrichtsausfällen.

Strukturierter Unterricht

Strukturierter Unterricht steht im Zusammenhang mit dem Element der kognitiven Unterstützung im Rahmen des Klassenklimas. Dazu gehört die Frage nach verständlichen Erklärungen des Lehrers oder der Lehrerin und der Hilfestellung, die Schülerinnen und Schüler bei Aufgaben bekommen. Bei strukturiertem Unterricht geht es grundsätzlich um häufige Interventionen zur Förderung des Lernprozesses, z. B. klare Benennung der Lernziele, Aufteilung des gesamten Lernstoffes in relativ kleine Einheiten, Präsentation dieser Einheiten in einer gut geplanten Abfolge, zahlreiche Übungsmöglichkeiten für die Schülerinnen und Schüler, Hilfestellung durch Hinweise, häufige Fragen und Überprüfungen, um den Lernfortschritt zu kontrollieren, und Feedback an die Schülerinnen und Schüler. Relevante Subkomponenten sind klare Zielsetzungen, Vorbereitung strukturierter Sequenzen von Unterrichts- und Lerneinheiten, verständliche Erklärungen, Einsatz von Fragen und Feedback als Teil der Wissensvermittlung und die Überprüfung der Schülerfortschritte.

Konstruktivistischer Unterricht und selbstbestimmtes Lernen

Konstruktivistischer Unterricht und selbstbestimmtes Lernen scheinen im Gegensatz zu den eher behavioristischen und gelenkten praktischen Ansätzen des strukturierten Unterrichts zu stehen. Konstruktivistischer Unterricht und selbstbestimmtes

Lernen unterstreichen den Lernprozess von Schülern und die Vermittlung allgemeiner und/oder fächergebundener Lernstrategien. „Kognitive Aktivierung“ ist ein wichtiges Element des konstruktivistischen Unterrichts. Dabei geht es um eine ausreichend tief gehende Präsentation des Inhalts, die auf ein Verstehen auf einer höheren Ebene abzielt, auf der Schülerinnen und Schüler verstehen, authentische Einsatzmöglichkeiten des Gelernten und Konzepte in unterschiedlichen Kontexten anwenden können. Die Idee vom „*Scaffolding*“, von unterstützenden Lerngerüsten, geht davon aus, dass das Maß an Selbstregulierung des Schülerlernens allmählich vergrößert wird, je besser die Schülerinnen und Schüler den Lernstoff beherrschen. Sie kombiniert die Konzepte von strukturiertem und konstruktivistischem Unterricht und stellt beide Ansätze in ein Kontinuum.

Differenzierung

Differenzierung erkennt die individuellen Unterschiede der Schülerinnen und Schüler an und versucht, Raum für Variationen im Unterricht zu schaffen, die diesen Unterschieden gerecht werden. Schulen und Lehrerinnen und Lehrer können den Unterricht dadurch differenziert gestalten, dass sie Schülerinnen und Schüler in Klassen kanalisieren, die auf unterschiedlichen Fähigkeitsniveaus arbeiten, Schülerinnen und Schüler innerhalb einer Klasse in Gruppen mit unterschiedlichen Fähigkeitsniveaus zusammenfassen, das Unterrichtstempo anpassen (und Schülern so die Möglichkeit geben, den Lernstoff nach unterschiedlichen Zeitplänen zu bearbeiten) und den Unterricht innerhalb relativ heterogener Klassenbesetzungen individualisieren. Zur Differenzierung gehören auch spezielle Programme und/oder zusätzlicher Unterricht und weitergehende Hilfen für schwache Schülerinnen und Schüler sowie Zusatzaufgaben für leistungsstarke Schülerinnen und Schüler.

8. Der Facettenreichtum schulwirksamkeitsfördernder Variablen

Es ist wichtig anzumerken, dass diese Konzepte sich nicht gegenseitig ausschließen und dass es mehrere Beispiele konzeptioneller Überschneidungen gibt. Leistungsorientierung wird nicht nur im Sinne direkter Schulphilosophien beschrieben, sondern auch als eine relevante Ausrichtung im Klima einer Schule. Außerdem ist es wichtig zu sehen, dass Evaluationen und die Überprüfung von Leistungen wesentliche Elemente der Leistungsorientierung einer Schule sind. Leistungsorientierung wirkt sich ebenso wie Schulleitungshandeln auf die Ergebnisse von Schülerleistungen aus. Abgesehen von der konzeptionellen Überschneidung dieser

Hauptindikatoren passiert es auch, dass Schulen eine Reihe dieser Indikatoren zu einer Gesamtstrategie in Bezug auf (eine Verbesserung von) Lehren und Lernen so kombinieren. Lerngelegenheiten und *Time-on-Task* sind z. B. wichtige Bestandteile einer umfassenden Strategie zur gesteigerten Konfrontation der Schülerinnen und Schüler mit Inhalten. „*Focused Teaching*“ (ein von Louis et al. 2010 geprägter Begriff) verbindet dagegen direkte Wissensvermittlung und konstruktivistische Unterrichtsstrategien. Ein drittes Beispiel ist die Art, wie die inhärente Verwandtschaft von Evaluation und Monitoring Lerngelegenheiten garantiert, was in Pophams Credo „*Test Better, Teach Better*“ (etwa: bessere Tests, besserer Unterricht; Popham 2003) seinen Niederschlag findet. Und schließlich kann man Evaluation und die Überprüfung von Leistungen als wirksame Hebel bei strukturierteren Unterrichtsansätzen betrachten, um Lernziele zu formulieren und den Schülerinnen und Schülern strukturierte Unterstützung und Anleitung zu geben. (Dieser Ansatz wird gerade in den Niederlanden unter der Überschrift „ergebnisorientierte Arbeit“ probeweise angewandt; vgl. Visscher & Ehren 2011.)

Viele der Faktoren wie Klima, Leistungsorientierung, Evaluation und Monitoring und Lerngelegenheiten, die oben erörtert wurden, lassen sich sowohl klassen- als auch schulbezogen messen und können auf beiden Ebenen unterschiedlich interpretiert werden. So mag ein Lehrer oder eine Lehrerin nur eingeschränkt von der Möglichkeit Gebrauch machen, die Überprüfung der Schülerleistungen in seinen Unterricht einfließen zu lassen, während der Schulleiter in der ganzen Schule ausgefeilte Systeme zur Kontrolle der Lernfortschritte der Schülerinnen und Schüler einsetzt, um schulbezogene Entwicklungsstrategien danach auszurichten.

Der Facettenreichtum der Schulwirksamkeit, die konzeptionellen Überschneidungen der Bedingungen und die zahlreichen Möglichkeiten, jede Subkomponente zu beschreiben und zu messen, machen deutlich, wie komplex die Gestaltung effektiver Inspektionskonzepte ist. Die Frage, welche Indikatoren Inspektionskonzepte enthalten sollen und wie sie zu messen sind, bedarf gründlicher Überlegungen. Im folgenden Abschnitt werden solche Überlegungen in Bezug auf sechs Inspektionskonzepte angestellt. Sie stammen aus den Niederlanden, England, Irland, Schweden, dem österreichischen Bundesland Steiermark und der Tschechischen Republik. Sie wurden von Ehren et al. (2013) im Rahmen eines von der EU finanzierten Projekts untersucht. Mithilfe der Dissertation von De Volder (2012) und der Länderprofile auf der Website des SICI, der *Standing International Conference of national and regional Inspectorates of education*, wurde die Tabelle vervollständigt. Die Beschreibung der Inspektionskonzepte bezieht sich auf das Jahr 2010.

9. Sechs Inspektionskonzepte: eine Analyse

Die folgende Übersicht bietet einen Überblick über die Inspektionskonzepte aus sechs europäischen Ländern. Die Länder repräsentieren eine große Bandbreite an Inspektoratstypen. Einige arbeiten nach einem *Low-Stakes*-Inspektionsansatz mit Betonung auf der Entwicklungskapazität einer Schule (z.B. Irland), andere überprüfen Schulen anhand testbasierter Frühwarn-Analysen (z.B. die Niederlande). Einige Länder haben stark zentralisierte nationale Inspektorate (z.B. England), andere Inspektorate arbeiten auf Regionalebene (z.B. in Österreich). In der Übersicht fassen wir die Inspektionsstandards und Schwellenwerte zusammen, nach denen die Inspektorate Schulen evaluieren. Darüber hinaus erläutern wir kurz den hierarchischen Aufbau dieser konzeptionellen Rahmen und die Art der Datensammlung, auf die sich die Beurteilung durch die Inspektorate stützt. Es ist wichtig anzumerken, dass sich hinter einer sichtbaren Überschneidung der Standards von Inspektionskonzepten sehr wahrscheinlich eine breitgefächerte Gamma-Verteilung operationaler Definitionen und spezifischer Messinstrumente verbirgt. Die Zusammenfassung und der nachfolgende Vergleich mit der Schulwirksamkeitsforschung muss sich daher auf allgemein gefasste Kommentare und Vorschläge beschränken, wie sich Inspektionskonzepte möglicherweise verbessern ließen.

Übersicht: Inspektionsstandards bei regelmäßig stattfindenden Gesamtinspektionen in sechs europäischen Ländern (Stand 2010)

Niederlande

Input

N/A

Prozess

- Die Schule verfügt über ein System, mit dem die Qualität ihres Unterrichts sichergestellt wird.
- Der den Schülerinnen und Schülern angebotene Lernstoff stellt eine ausreichende Vorbereitung auf die berufliche/akademische Bildung nach der weiterführenden Schule dar.
- Den Schülerinnen und Schülern steht ausreichend Unterrichtszeit zur Verfügung, um den Lernstoff zu bewältigen.
- Die Schule nimmt eine systematische Bewertung der Fortschritte der Schülerinnen und Schüler vor.
- Das Schulklima ist sicher und anregend.

- Das pädagogische Vorgehen der Lehrerinnen und Lehrer entspricht den grundlegenden Erfordernissen.
- Das didaktische Vorgehen der Lehrerinnen und Lehrer entspricht den grundlegenden Erfordernissen.
- Kinder mit sonderpädagogischem Förderbedarf bekommen die nötige Unterstützung.

Output

- Die Ergebnisse der Schülerinnen und Schüler entsprechen dem erwarteten Niveau (unter Berücksichtigung der Merkmale der Schülerpopulation).

Schwellen für „Failing Schools“

- Gesamtbeurteilung als sehr schwach, schwach, ausreichend oder gut.
- Eine Schule ist schwach, wenn die Leistungsergebnisse der Schülerinnen und Schüler am Ende der Grundschulzeit in mehr als drei aufeinanderfolgenden Jahren unzureichend ist.
- Eine Schule wird als sehr schwach beurteilt, wenn die Ergebnisse nach Abschluss der Grundschule in mehr als drei aufeinanderfolgenden Jahren unzureichend ist und das Lehren und Lernen oder die Betreuung der Schülerinnen und Schüler unzureichend ist.

Datensammlung und -auswertung

- Das Konzept beinhaltet die folgenden hierarchischen Ebenen: 5 Qualitätsbereiche, 10 Qualitätsaspekte, 46 Indikatoren (von denen 10 zum Bereich der Schwellen gehören).
- Es gibt Entscheidungsregeln für die Bewertung jedes einzelnen Indikators als ausreichend oder nicht ausreichend, außerdem Entscheidungsregeln für die Aggregation der Bewertungen der Indikatoren, um zu einer Beurteilung der Qualitätsaspekte, der Qualitätsbereiche und der Gesamtbewertung der Effektivität einer Schule zu gelangen.
- Datensammlung bei Inspektionsbesuchen im Rahmen von Unterrichtsbeobachtungen (mithilfe eines detaillierten Bewertungsbogens), leitfadengestützten Interviews mit Mitarbeitern der Schule, Eltern und Schülerinnen und Schülern und Dokumentenanalysen, u. a. Analyse der Selbstevaluation der Schule.
- Der Output der Schule wird mithilfe der Analyse landesweit standardisierter Tests und Abschlussprüfungen sowie anderer Indikatoren ermittelt (z. B. Abbrecher- und Abschlussquoten).

England

Input

Bericht zum Gebäudezustand, falls er Auswirkungen auf die Unterrichtsqualität oder Gesundheit und Sicherheit hat.

Prozess

- Qualität des schulischen Angebots, u. a. Unterricht, Lehrplan und Betreuung, Anleitung und Hilfestellung für Schülerinnen und Schüler
- Führung und Management
- Allgemeine Effektivität der Schule
- Fähigkeit der Schule zur Verbesserung

Output

- Leistung und Standards (Lernfortschritte der Schülerinnen und Schüler und in nationalen Tests erzielte Standards unter Berücksichtigung des Kontexts)
- Persönlichkeitsentwicklung und Wohlbefinden der Schülerinnen und Schüler

Schwellen für „Failing Schools“

- Schulen werden als *outstanding* (hervorragend), *good* (gut), *satisfactory* (zufriedenstellend) oder *inadequate* (ungenügend) eingestuft, wobei es bei den ungenügenden Schulen eine zusätzliche Unterteilung in „*Notice to Improve*“ (Aufforderung zur Verbesserung) und „*Special Measures*“ (besondere Maßnahmen) gibt. Bei etwa 40 Prozent der zufriedenstellenden Schulen wird der Bedarf beobachtender Inspektionen festgestellt.

Datensammlung und -auswertung

- Konzept beinhaltet 6 Hauptaspekte, unterteilt in 7 Kriterien zu Schülerergebnissen, 3 Kriterien zur Effektivität des schulischen Angebots und 8 Kriterien zur Effektivität von Führung und Management der Schulen.
- Datensammlung u. a. durch Unterrichtsbeobachtung (standardisierter Evaluationsbogen), leitfadengestützte Interviews mit Mitarbeitern der Schule, Eltern und Schülerinnen/Schülern und Dokumentenanalyse. Dazu gehört die Analyse der Selbstevaluation der Schule und die Auswertung von Fragebögen, die Eltern im Vorfeld des Inspektionsbesuchs ausfüllen.
- Messung des Outputs der Schule durch Analyse von Leistungsdaten: Lernfortschritte der Schülerinnen und Schüler, erreichte Standards in landesweiten Tests und Prüfungen unter Berücksichtigung von Kontextfaktoren (Anteil sozial benachteiligter Schülerinnen/Schüler, Anteil der Schülerinnen/Schüler, deren Muttersprache nicht Englisch ist), Analyse schriftlicher Schülerarbeiten und durch die Schule zur Verfügung gestellter Leistungsdaten.

Schweden

Input

N/A

Prozess

- Der Fokus liegt auf der Einhaltung gesetzlicher Vorschriften, mit Inspektionen soll sichergestellt werden, dass die Rechte der Schülerinnen und Schüler in Bezug auf Bildung berücksichtigt werden.
- Ergebnisse, Leistungsstandards, Lernen und Lehren: wie Schulen auf eine Verbesserung der Leistungsstandards hinarbeiten, wie Lehrerinnen und Lehrer ihren Unterricht auf die individuellen Bedürfnisse der Schülerinnen und Schüler ausrichten, ob und wie sie Erkenntnisse aus den Schülerergebnissen in ihrem Unterricht umsetzen, wie und ob sie Kinder mit speziellem Förderbedarf unterstützen.
- Den Schülerinnen und Schülern die Regeln und Werte einer demokratischen Gesellschaft nahebringen.
- Management und internes Audit-Verfahren: Schulverwaltung und -management.

Output

- Ergebnisse, Leistungsstandards, Lernen und Lehren: was Schülerinnen und Schüler mit Blick auf die im nationalen Curriculum formulierten Ziele lernen.

Schwellen für „Failing Schools“

- Leitfaden für Beurteilungen wird erstellt, zur Zeit ganzheitlicher Ansatz, bei dem man sich bei der Einschätzung, ob sich eine Schule an die gesetzlichen Vorgaben hält, auf Erfahrung und Kompetenz des Inspektors verlässt.

Datensammlung und -auswertung

- Konzept beinhaltet 3 Hauptbereiche und 16 Indikatoren.
- Datensammlung u. a. durch Unterrichtsbeobachtung (Anmerkungen zu festgelegten Punkten auf einem Beurteilungsbogen, keine detaillierte Punktwertung), Befragung des Direktoriums, Dokumentenanalyse (Statistiken, Qualitätsberichte, vorangegangene Kontrollbeschlüsse und Inspektionsberichte, Selbstevaluationen), Interviews mit Schulpersonal, u. a. Schulleitung, Lehrerinnen/Lehrern, Schulkrankenschwestern und Schülerinnen/Schülern, Eltern und zuständigen Politikern.

Irland

Input

- Informationen zur Ressourcenverteilung für Schülerinnen und Schüler mit besonderem Förderbedarf, außerdem Informationen zu Bauprogrammen und Plänen zur Verbesserung oder Erweiterung der Einrichtungen (keine separate Überprüfung).

Prozess

Auf Grundschulniveau:

- Qualität des Schulmanagements
- Qualität der Schulplanung
- Qualität von Unterricht und Lernen
- Qualität der Unterstützung der Schülerinnen und Schüler

Output

- Keine Output-Standards. Testergebnisse werden während des Inspektionsbesuchs analysiert, jedoch nicht in den Bericht aufgenommen. Das Analyseergebnis von Testnoten kann mit der Schulleitung und/oder Lehrerinnen und Lehrern besprochen werden und trägt zum Urteil der Inspektoren über die Unterrichts- und Lernqualität bei. Statistische Daten fließen jedoch nicht in den Inspektionsbericht ein.

Schwellen für „Failing Schools“

- Keine Schwellen. Gegenwärtig ist die Evaluation nicht Grundlage für eine Gesamtwertung. Empfehlungen für künftige Maßnahmen des Schulmanagements oder Fachlehrerinnen und Fachlehrer werden formuliert.

Datensammlung und –auswertung

- Konzept umfasst 5 Evaluationsbereiche, unterteilt in 143 „Themen der Selbstevaluation“. Theoretisch müssen Schulen Nachweismaterial sammeln und sich dann selbst mit ihrer Leistung zu jedem Thema auf einer vierteiligen Skala einordnen. Dieser Selbstevaluationsprozess bildet die Grundlage eines Inspektionsteams, das Gesamtevaluationen von Schulen (*Whole School Evaluations*) durchführt.
- Datensammlung mittels Dokumentenanalyse (Formular mit Schulinformationen, von Schulleitung ausgefüllt, Dokumentation der Schule zu Schüleranmeldungen, Vorgaben des Lehrplans, Planungen, Schulrichtlinien usw.), Fragebögen an Schulleitung und Fachbereichskoordinatoren, Interviews mit Mitarbeitern der Schulen, Unterrichtsbeobachtungen.

- Output-Daten werden gesammelt, auf Grundschulniveau vergleichen die Inspektorinnen und Inspektoren die Schülerleistungen in standardisierten Tests in Mathematik und Englisch mit nationalen Normen.
- Auf Grundschulniveau analysieren die Inspektorinnen und Inspektoren laufende Tests, die Lehrerinnen und Lehrer in allen Fächern einsetzen, sowohl für formative als auch für summative Zwecke, außerdem standardisierte, von einer externen Forschungsagentur entwickelte Tests in Englisch und Mathematik, die die Lehrerinnen und Lehrer der Schule kaufen und stellen.

Österreich (Bundesland Steiermark)

Input

- N/A

Prozess

- Lehren und Lernen
- Klassenzimmer und Schulen als Lebensraum
- Schulpartnerschaft und Außenbeziehungen
- Schulmanagement
- Professionalität und Personalentwicklung

Output

- N/A
- Es gibt standardbasierte Leitungstests, doch Inspektorate dürfen sie nicht einsetzen oder in die Output-Messungen einbeziehen.

Schwellen für „Failing Schools“

- Keine Schwellen, ein Überblick über Stärken und Schwächen wird formuliert. Über Konsequenzen und notwendige Maßnahmen werden die Schule und die zuständigen Behörden informiert.

Datensammlung und -auswertung

- Konzept umfasst 5 Bereiche der Schulqualität, unterteilt in 16 Kriterien, mit einem Fokus auf einer kritischen Analyse der Selbstevaluationsprozedur und -ergebnisse der Schulen, wie sie im Schulprogramm formuliert ist. Nur wenn diese Dokumentenanalyse („methodisch-substantive Analyse der Selbstevaluation“) Mängel aufweist (in Bezug auf die Prozeduren der Selbstevaluation) oder aufzeigt, dass „wesentliche Anforderungen und Standards“ (bezieht sich auf die Ergebnisse der Selbstevaluation) nicht eingehalten werden, muss eine

„spezifische Untersuchung“ durchgeführt werden. Einer solchen Untersuchung gehen Beratungen und Hilfsmaßnahmen voraus.

- Datensammlung: Dokumentenanalyse des „Inspektionsordners“ (von der Schule zusammengestellt), in dem alle Schulentwicklungspläne der letzten 5 Jahre enthalten sind, außerdem das Schulprogramm für die individuelle Förderung aller Schülerinnen und Schüler, eine Liste aller Lehrer (mit Altersangabe), Stundenpläne, Liste der Schülerinnen und Schüler, Unterrichtsbeobachtungen (strukturiertes Protokoll, das auch den Lehrerinnen und Lehrern zur Verfügung steht), Interviews mit Lehrerinnen/Lehrern, Schülerinnen/Schülern, Eltern, Schulleitung und Bürgermeister (mithilfe eines strukturierten Leitfadens).

Tschechische Republik

Input

- Bei einer „Zustandsüberprüfung“ kontrollieren die Prüfer, ob gesetzliche Bestimmungen zur Bereitstellung schulischer Angebote erfüllt sind (zusätzlich zur Prüfung der Gebäude).
- Arbeitsbedingungen Personal
- Materielle Voraussetzungen
- Finanzielle Voraussetzungen

Prozess

- Gleiche Bildungschancen
- Bildungsprogramme der Schulen
- Schulmanagement
- Effektive Schulorganisation
- Effektive Unterstützung der Persönlichkeitsentwicklung der Kinder
- Partnerschaft
- Effektive Unterstützung der Entwicklung von Schlüsselkompetenzen von Kindern, Schülern und Studenten

Output

- Systematische Evaluation einzelner Schülerinnen/Schüler- und Gruppenergebnisse
- Evaluation von Gesamtergebnissen einer Schule

Schwellen für „Failing Schools“

- Das Inspektorat evaluiert die Schulen mithilfe einer vierstufigen Skala: A, B, C, D. Das Evaluationsergebnis A kann zu einer Streichung der Schule aus dem Schulregister führen. Das Evaluationsergebnis D wird als vorbildlich angesehen.

Datensammlung und -auswertung

- Konzept umfasst 7 Hauptbereiche, aufgeteilt in 22 Indikatoren, die in 60 Subindikatoren unterteilt sind.
- Datensammlung: Unterrichtsbeobachtung (mit Beobachtungsprotokoll), Analyse der Ergebnisse der Selbstevaluation der Schulen, Inspektionsanalyse der Schuldokumente, Inspektionen vor Ort, Beobachtungen der Schülerinnen/Schüler und Lehrerinnen/Lehrer im Unterricht und standardisierte Interviews mit pädagogischen Mitarbeitern sowie Kontakt zu Schulgründern.
- Output wird anhand der Abschlussprüfungen der Oberstufen weiterführender Schulen gemessen, die wesentlich in der Verantwortung der Einzelschulen liegen. Keine nationalen Prüfungen, keine Tests in Grundschulen.
- Standardisierte Tests seit 2011

Die Zusammenfassung in der Übersicht zeigt, dass sich alle sechs Inspektorate auf beeinflussbare Bedingungen und Prozesse konzentrieren, einige jedoch zusätzlich noch den Output der Schule evaluieren. Und nur eines der Inspektorate befasst sich ausdrücklich mit Aspekten des Inputs von Schulen (in der Tschechischen Republik). Interessanterweise sammeln zwei Länder (Irland und England) Informationen zum Input (z. B. die Qualität der Schulgebäude), verwenden diese Informationen jedoch nur, wenn die Inspektorinnen und Inspektoren der Ansicht sind, dass dieser Input sich besonders auf die Qualität der schulischen Prozesse auswirkt. Diese Inspektorate benoten den Input einer Schule nicht separat. Die Wahl dieses Ansatzes scheint die komplexe Verknüpfung der Input- und Prozessbedingungen mit Schulqualität und guten Schülerleistungen anzuerkennen.

Eine solche Sichtweise fehlt jedoch in der generellen Bewertung von Schulprozessen und Schul-Output durch die sechs Inspektorate. Die Übersicht zeigt auf, dass sich die Gesamtbewertung der Schulqualität auf Standards, Substandards und weitere Unterteilungen in detaillierte Kriterien stützt, wobei bei der Benotung von Schulen auf einer Skala strenge Regeln in ähnlicher Weise auf alle Schulen angewandt werden. Bei den Beurteilungsprotokollen und Leitfäden muss eine Bewertung der Bedingungen mit „vorhanden“ oder „nicht vorhanden“ oder eine Bewertung auf einer 3- oder 4-stufigen Skala angegeben werden. Dabei wird die mögliche Verbindung der Bedingungen untereinander nicht berücksichtigt.

Die Art und Weise, wie Informationen auf Klassenebene aggregiert werden, um die Wirksamkeit auf Schulebene zu evaluieren, lässt ebenfalls die Tatsache unberücksichtigt, dass diese Bedingungen auf unterschiedlichen Ebenen in der Schulhierarchie unterschiedliche Bedeutungen haben. Feststellungen zur „Leistungsorientierung“ bei Unterrichtsbeobachtungen (auf Klassenebene) werden z. B. oft einfach gemittelt, um zu einer Bewertung der Qualität der Schule in diesem Bereich zu kommen. Auf Schulebene kann Leistungsorientierung jedoch auch eine Bewertung

der Schulphilosophie in diesem Bereich beinhalten und Kontrollsysteme einschließen, die die Schule aufgebaut hat, um die Ausrichtung der Lehrerinnen und Lehrer auf gute Schülerleistungen zu unterstützen. Es sind nur einige der Schulinspektorate (z. B. Schweden und Österreich), die einen Bogen um diesen Ansatz machen, der Schul- und Klassenbedingungen über einen Kamm schert. Diese Schulinspektorate haben eine ganzheitlichere Herangehensweise an Schulevaluationen und geben Schulen einen Überblick über ihre Stärken und Schwächen, statt mit rudimentären Schwellen zu arbeiten, um „*Failing Schools*“ herauszugreifen. Der Nachteil bei diesem Ansatz besteht jedoch in einem möglichen Mangel an Genauigkeit und Transparenz in den Inspektionsbewertungen.

Zum Thema des pauschalen oder spezifischer operationalisierten Inspektionsstandards bleibt noch anzumerken, dass die Beispiele aus quantitativen Indikatorensystemen, wie sie in diesem Kapitel angeführt werden, in den meisten Fällen spezifischer sind als die Inspektionsstandards, die in Tabelle 1 gelistet sind. Dabei sollte man jedoch die Tatsache nicht aus den Augen verlieren, dass sich systematische quantitative Indikatorensysteme von Schulinspektionen insofern unterscheiden, als Letztere den großen Vorteil des professionellen Urteilsvermögens von Experten mit sich bringen, während Erstere von Messtechniken abhängen. Trotzdem könnte man die extensiveren Indikatorensets als Ressource bei der Entwicklung von Inspektionskonzepten verwenden und möglicherweise ließen sie sich teilweise als Grundlage für strukturierte operative Abläufe bei Klassenbeobachtungen und als „Stützgerüste“ für die professionellen Einschätzungen von Schulinspektoren und -inspektoren übernehmen. Was Ausrichtung und Gestaltung von Inspektionskonzepten angeht, so gibt es viele Optionen und eine große Auswahl an Instrumenten. Das von Scheerens et al. (2011) entwickelte Indikatorenset, das im Anhang vorliegt, soll als Beispiel für eine sachbezogene Zusammenstellung von Inspektionsstandards dienen.

Einige Schulinspektorate (z. B. OFSTED in England) haben solche detaillierten Checklisten jedoch aufgegeben, nachdem standardisierte und schriftlich festgelegte Rezepte zur Schulentwicklung aus ihnen geworden waren. Schulen und andere Stakeholder (z. B. Partner bei der Schulverbesserung und Entwickler von Selbstevaluationen für Schulen) entwickelten sogenannte „OFSTED-genehmigte“ Modelle für Schulorganisation und Unterricht, die immer öfter von Schulen kopiert und eingeführt wurden, ohne dass geprüft wurde, ob diese Praktiken für die spezielle Klasse oder den Schulkontext geeignet waren, auf den sie angewendet werden sollten.

Die Übersicht über die Inspektionskonzepte fasst auch zusammen, wie eine Reihe von Schulinspektoren den Output von Schulen in ihre Evaluation der Schulqualität einbeziehen. Die Beschreibung, wie Testdaten analysiert werden, um den Schul-Output zu bewerten, macht deutlich, dass diese Evaluationen immer noch recht rudimentär sind und nur in eingeschränktem Maße auf die zu Be-

ginn des Kapitels dargestellten ausgefeilteren *Value-Added*-Modelle für Analyse und Bericht von Schul-Output zurückgreifen. Die Schulinspektorate, die Zugang zu Schülerleistungsdaten haben (in den Niederlanden und England) berücksichtigen eine mögliche Instabilität in den Daten, indem sie bei der Bewertung des Schul-Outputs Durchschnittswerte berechnen (normalerweise aus drei Jahren). Die meisten Schulinspektorate scheinen jedoch keine detaillierten und qualitativ hochwertigen Leistungsdaten zur Verfügung zu haben, um solche Analysen vornehmen zu können. Entweder gibt es keine nationalen standardisierten Tests oder sie dürfen nicht auf solche Daten zugreifen. Stattdessen analysieren einige Schulinspektorate Bewertungen durch Lehrerinnen und Lehrer oder Schülerarbeiten, um sich eine Vorstellung vom Output einer Schule zu machen. Interessanterweise werden solche Analysen in manchen Fällen (z.B. in Irland) auch benutzt, um die Qualität von Prozessen zu beurteilen (z.B. die Unterrichtsqualität). Dadurch dienen sie als stellvertretende Messgröße für die Qualität bildungsbezogener Prozesse an der Schule.

An der Übersicht kann man die bemerkenswerte Tatsache ablesen, dass die Unterrichtsqualität zwar als Hauptbedingung für Schulqualität angesehen und entsprechend betont wird, dass aber keines der Schulinspektorate die Merkmale von Lehrerinnen und Lehrern in sein Konzept der Inspektionsstandards aufgenommen hat. Obwohl das niederländische Inspektionskonzept Kriterien auf Klassenebene wie das pädagogische und didaktische Verhalten von Lehrerinnen und Lehrern einbezieht, werden sie lediglich auf Schulebene bewertet, wenn die Ergebnisse einer Auswahl von Unterrichtsbeobachtungen gemittelt werden. Wie wir bereits ausgeführt haben, sind jedoch Bedingungen auf Unterrichtsebene wie hohe Erwartungen, ein anspruchsvoller Unterrichtsansatz, eine geordnete Lernumgebung und klarer und strukturierter Unterricht wichtiger für die Verbesserung der Schülerleistungen als Bedingungen auf Schulebene. Die meisten Inspektorate evaluieren Lehrerinnen und Lehrer oder Unterricht jedoch nicht ausdrücklich auf der Ebene von Klasse/Fach oder Klassenstufe, sondern evaluieren lieber die Bedingungen auf Schulebene und allgemeine Merkmale des Unterrichts oder Unterrichtsmuster wie Lernzeit, Schulleitungshandeln und Schulklima.

Eine abschließende „Empfehlung“ für zukünftige und weiterreichende Entwicklungen nationaler Schulinspektionskonzepte ist daher, die Betonung von Standards schulischer Organisation in Relation zu Lehre und Lernen oder didaktischen Standards zu setzen. Eine vorsichtige Warnung an Gestalter von Schulinspektionskonzepten könnte sein, die primären Prozesse des Lehrens und Lernens nicht aus den Augen zu verlieren.

Literatuur

- American Statistical Association (2014). *ASA statement on using value-added models for educational assessment*. Alexandria, VA: Author. Verfügbar unter: https://www.amstat.org/policy/pdfs/ASA_VAM_Statement.pdf. [Zugriff am 17.03.2015]
- Bosker, R. J., Guldemon, H. G., Hofman, R. H. & Hofman, W. H. A. (1989). De stabiliteit van schoolkwaliteit. In J. Scheerens & J. C. Verhoeven (Hg.), *Schoolorganisatie, beleid en onderwijskwaliteit* (S. 39–53). Lisse: Swets & Zeitlinger.
- Cotton, K. (1995). *Effective schooling practices: A research synthesis. 1995 Update* (School Improvement Research Series). Northwest Regional Educational Laboratory.
- Creemers, B. P. M. & Kyriakides, L. (2008). *The Dynamics of Educational Effectiveness*. London, New York: Routledge.
- De Grauwe, A. (2007). *Module 7. Alternative models in reforming school supervision*. Verfügbar unter: http://www.unesco.org/iiep/PDF/TR_Mods/SUP_Mod7.pdf [Zugriff am 17.03.2015]
- De Volder, I. (2012). *Externe schoolevaluaties in Europa. Een vergelijkend onderzoek*. Antwerpen: Grant uitgevers.
- Dijkstra, A. B., De la Motte, P. I., Ehren, M. C. M. & Eilard, A. (2014). Discussion. School inspections and school improvement in the social domain. The assessment of social outcomes of education. In: A. B. Dijkstra & P. I. de la Motte (Hg.). *Social Outcomes of Education; The assessment of social outcomes and school improvement through school inspections* (S. 189–215). Amsterdam: Amsterdam University Press.
- Donaldson, M. L. & Johnson, S. M. (2010). The price of misassignment: The role of teaching assignments in Teach for America teachers' exit from low-income schools and the teaching profession. *Educational Evaluation and Policy Analysis*, 32 (2), 299–323.
- Eddy Spicer, D., Ehren, M., Bangpan, M., Khatwa, M. (2014). *Under what conditions do inspection, monitoring and assessment improve system efficiency, service delivery and learning outcomes for the poorest and most marginalised? A realist synthesis of school accountability in low- and middle-income countries. Protocol*. London: EPPI-Centre, Social Science Research Centre, Institute of Education, University of London. Verfügbar unter: <http://eppi.ioe.ac.uk/cms/LinkClick.aspx?fileticket=x-3hzguBXhY%3D&tabid=3174> [Zugriff am 17.03.2015]
- Ehren, M. C. M., Altrichter, H., McNamara, G. & O'Hara, J. (2013). Impact of school inspections on school improvement; describing assumptions on causal mechanisms in six European countries. *Educational Assessment, Evaluation and Accountability*, 25 (1), 3–43. Verfügbar unter: <http://dx.doi.org/10.1007/s11092-012-9156-4> [Zugriff am 17.03.2015]
- Ehren, M. C. M. & Dijkstra, A. B. (2014). Evaluating Social Outcomes through School Inspections. In: A. B. Dijkstra & P. I. de la Motte (Hg.). *Social Outcomes of Education; The assessment of social outcomes and school improvement through school inspections* (S. 51–73). Amsterdam: Amsterdam University Press.

- Gray, J., Jesson, D., Goldstein, H., Hedges, K. & Rasbash, J. (1995). A multi-level analysis of school improvement: Changes in school's performance over time. *School Effectiveness and School Improvement*, 6, 97–114.
- Hamilton, L. S. & Koretz, D. M. (2002). Tests and their use in test-based accountability systems. In L. S. Hamilton, B. M. Stecher & S. P. Klein (Hg.). *Making sense of Test-based Accountability in Education*. Santa Monica: Rand cooperation. Verfügbar unter: http://www.rand.org/pubs/monograph_reports/MR1554/ [Zugriff am 17.03.2015]
- Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: an update. *Educational Evaluation and Policy Analysis*, 19 (2), 141–164.
- Hanushek, E. A. (1986). The Economics of Schooling. Production and Efficiency in Public Schools. *Journal of Economic Literature*, 24 (3), 1141–1177
- Hattie, J. (2009). *Visible Learning*. Abingdon: Routledge.
- Hopkins, D., Stringfield, S., Harries, A., Stoll, L. & Mackay, T. (2014). School and System Improvement. *School Effectiveness and School Improvement*, 25, 257–281.
- Levine, D. K. & Lezotte, L. W. (1990). *Unusually Effective Schools: A Review and Analysis of Research and Practice*. Madison, Wise: Nat. Centre for Effective Schools Research and Development.
- Louis, K. S., Dretzke, B. & Wahlstrom, K. (2010). How does leadership affect student achievement? Results from a national US survey. *School Effectiveness and School Improvement*, 21 (3), 315–336.
- Marzano, R. (2003). *What works in schools. Translating research into action*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Mourshed, M., Chijioke, C. & Barber, M. (2010). *How the world's most improved school systems keep getting better*. London: McKinsey. Verfügbar unter: <http://mckinseyon-society.com/how-the-worlds-most-improved-school-systems-keep-getting-better/> [Zugriff am 17.03.2015]
- Muijs, D., Creemers, B., Kyriakides, L., Van der Werf, G., Timperley, H. & Earl, L. (2014). Teaching Effectiveness. A State of the Art Review. *School Effectiveness and School Improvement*, 24, 231–256.
- Popham, W. J. (2003). *Test better, teach better. The instructional role of assessment*. Alexandria, VA: USAL ASCD.
- Purkey, S. C. & Smith, M. S. (1983). Effective schools: a review. *The Elementary School Journal*, 83, 427–452.
- Reynolds, D., Sammons, P., De Fraine, B., Townsend, T., Van Damme, J., Teddlie, C. & Stringfield, S. (2014). Educational Effectiveness Research (EER): A state of the Art Review. *School Effectiveness and School Improvement*, 25 (2), 197–230.
- Sammons, P., Hillman, J. & Mortimore, P. (1995). *Key characteristics of effective schools: A review of school effectiveness research*. London: OFSTED.
- Scheerens, J. (1992). *Effective Schooling, Research, Theory and Practice*. London: Cassell.
- Scheerens, J. (2007) *Conceptual framework for the PISA 2009 background questionnaires*. Internal report, University of Twente.

- Scheerens, J. (2013). *What is effective schooling? A review of current thought and practice*. Paper for the International Baccalaureate Organization; Washington DC.
- Scheerens, J. (2014) School, teaching and system effectiveness: some comments on three state of the art reviews. *School effectiveness and School Improvement*, 25, 282–290.
- Scheerens, J. & Creemers, B. P. M. (1989). Conceptualizing school effectiveness. Development in school effectiveness research. *Special issue of the International Journal of Educational Research*, 13 (7), 691–706.
- Scheerens, J., Glas, C. A. W. & Thomas, S. M. (2003). *Educational evaluation, assessment, and monitoring: a systemic approach*. Lisse: Swets & Zeitlinger.
- Scheerens, J., Luyten, H., Van den Bergh, S. M. & Glas C. A. W. (2015). Exploration of direct and indirect associations of system level policy amenable variables and reading literacy performance. *Educational Research and Evaluation*. (im Erscheinen)
- Scheerens, J., Luyten, H., Steen, R. & Luyten-de Thouars, Y. (2007). *Review and meta-analyses of school and teaching effectiveness*. Enschede: University of Twente, Department of Educational Organisation and Management.
- Seidel, T. & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: the role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77 (4), 454–499.
- Thomas, S. M., Peng, W. J. & Gray, J. (2010). Modeling patterns of improvement over time: value added trends in English secondary schools. *Oxford Review of Education*, 33, 261–295.
- Tolofari, S. (2005) „New public management and education“. *Policy Futures in Education*, 3 (1), 75–89. Verfügbar unter: <http://firgoa.usc.es/drupal/files/tolofari.pdf> [Zugriff am 17.03.2015]
- Van Bruggen, J. C. (2010). *Inspectorates of Education in Europe: some comparative remarks about their tasks and work*. Standing International Conference of Inspectorates of Education in Europe (SICI). Verfügbar unter: <http://www.sici-inspectorates.eu/Members/Comparative-Analyses/Inspectorates-of-Education-in-Europe> [Zugriff 17.03.2015]
- Vermeer, N. & Van der Steeg, M. (2011). *Onderwijsprestaties Nederland in Internationaal Perspectief*. CPB Achtergronddocument bij CPB Policy Brief 05, 2011. Den Haag: CPB.
- Visscher, A. & Ehren, M. (2011). *De eenvoud en complexiteit van Opbrengstgericht Werken*. Verfügbar unter: <http://www.rijksoverheid.nl/documenten-en-publicaties/rapporten/2011/07/13/de-eenvoud-en-complexiteit-van-opbrengstgericht-werken.html> [Zugriff 30.03.2015]

Anhang

Detaillierteres Indikatorenset für Output, Input und schulische Prozesse
(Scheerens et al. 2011)

Tabelle 1: Überblick über Indikatoren für Bildungs-Outcome

Hauptkategorien von Outcome-Indikatoren	Unterkategorien	Technische Angaben
Output-Indikatoren	Leistungsmessungen fächergestützt Literacy (Lesen, Mathematik, Naturwissenschaft) Kompetenzen (z. B. das Lernen lernen)	Effektmessung des <i>Value Added</i> ; Wachstumskurven Bewertungsmethodik (von Multiple-Choice-Tests bis hin zu authentischen Leistungsmes- sungen) Kriteriums- bzw. norm- orientierte Testverfahren
Indikatoren für Out- come/Zielerreichung	Messungen der Zielerreichung Schulabschlussquote Anteil der Schülerinnen und Schüler, die ihren Schulab- schluss ohne Wiederholung einer Klasse erlangt haben Quote der Abbrecher Quote der Wiederholer	Auswahlorientierte Schulricht- linien, statistisch kontrolliert

Tabelle 2: Indikatoren für finanzielle und materielle Ausstattung auf Schulebene

Finanzielle und materielle Ausstattung auf Schulebene
<ul style="list-style-type: none"> • Anteil des Schulbudgets, der nicht aus öffentlichen Quellen stammt • Schulgebäude und Ausstattung • Ausstattung der Klassenräume (Möbiliar, Computer usw.) • Schulartikel wie Bleistifte und Papier, Schultafel, Flipchart • Verfügbarkeit von Lehrbüchern in den wichtigsten Unterrichtsfächern • Grundausrüstung wie getrennte Toiletten für Jungen und Mädchen, Wasser, Strom, Heizung, Telefon, Verfügbarkeit von zusätzlichen Leistungen in Bezug auf Ernährung, Gesundheit und Beförderung

Tabelle 3: Überblick über Beispiele für Prozessindikatoren für Funktionsweisen von Schulen

Prozessindikatoren auf Schulebene
<p><i>Einbeziehung der Umgebung</i></p> <ul style="list-style-type: none"> • Der Grad der tatsächlichen Einbeziehung von Eltern in unterschiedliche Schulaktivitäten (Lehr- und Lernprozess, außerschulische Aktivitäten, unterstützende Aktivitäten) • Der prozentuale Anteil des gesamten jährlichen Schulhaushaltes, der aus Quellen vor Ort finanziert wird • Das Maß der Mitsprachemöglichkeiten, die Schulträger in Bezug auf die Arbeitsbedingungen von Lehrerinnen und Lehrern haben [mögliche Operationalisierungen im EDUCO-Projekt, El Salvador]
<p><i>Finanzielle und personelle Mittel der Schule</i></p> <ul style="list-style-type: none"> • durchschnittliche Berufserfahrung der Lehrerinnen und Lehrer einer Schule in Jahren • zahlenmäßiges Schüler-Lehrer-Verhältnis an einer Schule • durchschnittliche Klassenstärke pro Schule • Anteil der Lehrerinnen und Lehrer mit regulärer Ausbildung pro Schule • Personelle Ausstattung der „Spitze“ einer Schule (Schulleiter und stellvertretender Schulleiter pro 100 Schülerinnen und Schüler)
<p><i>Leistungsorientierte Ausrichtung</i></p> <ul style="list-style-type: none"> • Ob Schulen Leistungsstandards formulieren oder nicht • Wie intensiv Schulen den (bildungsbezogenen) Werdegang ihrer Schülerinnen und Schüler verfolgen, nachdem sie die Schule verlassen haben • Ob Schulen Leistungsergebnisse an interessierte Zielgruppen weitergeben oder nicht
<p><i>Schulleitungshandeln</i></p> <ul style="list-style-type: none"> • Zeit, die Schulleiter auf pädagogische Arbeit verwenden (im Vergleich zu Verwaltungsangelegenheiten oder anderen Aufgaben) • Ob Schulleiter die Leistung von Lehrerinnen und Lehrern bewerten oder nicht • Zeit, die bei Konferenzen auf unterrichtsbezogene Themen verwendet wird
<p><i>Kontinuität und Konsens im Kollegium</i></p> <ul style="list-style-type: none"> • Ausmaß personeller Veränderungen in einem bestimmten Zeitraum • Existenz oder Fehlen fächerbezogener Arbeitsgruppen oder Abteilungen (weiterführende Schulen) • Häufigkeit und Dauer formeller und informeller Zusammenkünfte des Kollegiums
<p><i>Geregeltes und sicheres Klima</i></p> <ul style="list-style-type: none"> • Statistisches Erfassen von häufigem Fehlen oder Schulschwänzen • Beurteilung der Schuldisziplin durch Schulleiter, Lehrerinnen/Lehrer und Schülerinnen/Schüler

Effiziente Nutzung der Zeit

- Gesamte Unterrichtszeit und Zeit pro Fach
- Durchschnittlicher Zeitverlust pro Unterrichtsstunde (durch organisatorische Notwendigkeiten, Raum- oder Gebäudewechsel, Störungen)
- Prozentualer Anteil ausgefallener Stunden pro Jahr

Lerngelegenheiten

- Beurteilung durch Lehrerinnen und Lehrer oder Schülerinnen und Schüler, ob jedes Item in einem Leistungstest unterrichtet wurde oder nicht

Evaluation der Lernfortschritte der Schülerinnen und Schüler

- Häufigkeit von curriculumspezifischen Test in den einzelnen Klassenstufen
- Häufigkeit von standardisierten Leistungstests
- Testergebnisse und ihre Verwendung durch die Lehrerinnen und Lehrer

Bewertung der Lehrerqualität

- Beurteilung der Unterrichtsqualität durch Peers (andere Lehrerinnen und Lehrer)
- Beurteilung der Unterrichtsqualität durch Schülerinnen und Schüler

Tabelle 4: Überblick über effektive Unterrichts- und Lernvariablen

Effektive Unterrichtsvariablen
<p><i>Hauptsächliche Unterrichtsfaktoren</i></p> <ul style="list-style-type: none"> • Lerngelegenheiten • Struktur und stützende Gerüste (kognitive Struktur) • Anregende Einbeziehung (Motivationsstruktur) • Aspekte des Klimas: <ul style="list-style-type: none"> • aufgabenorientiert • gegenseitiger Respekt • geordnet, sicher • Monitoring und Fragen • Feedback und Bestätigung • Modellhaftes Vermitteln von Lern- und Selbstregulierungsstrategien • „Authentische“ Anwendungen • Adaptiver Unterricht
<p><i>Lernstrategien von Schülerinnen und Schülern</i></p> <ul style="list-style-type: none"> • sichtbar: <ul style="list-style-type: none"> • Aktive Lernzeit • Arbeit mit Materialien • Kooperatives Lernen • mental: <ul style="list-style-type: none"> • Selbstregulierung • Selbstkontrolle • Meta-kognitive „Handlungen“ • Lernstile